

R E P O R T R E S U M E S

ED 003 878

48

COMPARING THE PHONETIC CHARACTERISTICS OF LANGUAGES.

BY- DELATTRE, PIERRE

COLORADO UNIV., BOULDER

REPORT NUMBER NDEA-VI-70

PUB DATE 15 JUN 63

CONTRACT OEC-2-14-026

EDRS PRICE MF-\$0.27 HC-\$6.16 154P.

DESCRIPTORS- #PHONETIC ANALYSIS, #LANGUAGE RESEARCH,
#COMPARATIVE ANALYSIS, #LANGUAGE PATTERNS, FRENCH, GERMAN,
SPANISH, ENGLISH, #RESEARCH METHODOLOGY, BOULDER, COLORADO

THE PHONETIC FEATURES NECESSARY FOR GOOD PRONUNCIATION AND EFFECTIVE USE OF LANGUAGES WERE STUDIED. A FOUR-WAY RESEARCH TECHNIQUE WAS DEVELOPED FOR COMPARISON AMONG LANGUAGES (1) INVOLVING SPECTROGRAPHIC ANALYSIS, (2) SPECTROGRAPHIC SYNTHESIS, (3) MOTION PICTURE ANALYSIS, AND (4) STATISTICAL ANALYSIS. THE ELECTRONIC INSTRUMENTATION CONSTRUCTED FOR THE STUDY WAS ALSO DESCRIBED. THE PROSODIC, VOCALIC, AND CONSONANT FEATURES OF FRENCH, ENGLISH, GERMAN, AND SPANISH COMPARED UNDER 40 SUBDIVISIONS WERE DISCUSSED IN DETAIL. SOME SIGNIFICANT RESULTS OF THE STUDY WERE (1) AN X-RAY MOTION PICTURE LIBRARY OF TONGUE PROFILES FOR FRENCH AND ENGLISH, (2) AN ISOLATION OF CONSONANT VOICING FACTORS BY SYNTHESIS FOR ENGLISH, AND (3) VARIOUS SUCCESSFUL PHONETIC COMPARISONS. RELATED MATERIAL MAY BE FOUND IN ED 003 879 AND ED 010 231. (RS)

ED0003878

70

COMPARING THE PHONETIC CHARACTERISTICS OF LANGUAGES

Pierre Delattre

University of Colorado

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education

This document has been reproduced exactly as received from the
person or organization originating it. Points of view or opinions
stated do not necessarily represent official Office of Education
position or policy.

DEPARTMENT OF FRENCH
UNIVERSITY OF COLORADO
Boulder, Colorado

Comparing the Phonetic Characteristics of Languages

by

Pierre Delattre

Office of Education Contract OE-2-14-026

Final Report

15 June 1963

(Second Interim Report)

Pierre Delattre

Pierre Delattre
Project Director

The research reported herein was
performed pursuant to a contract
with the United States Office of
Education, Department of Health,
Education, and Welfare.

TABLE OF CONTENTS

PROSODIC FEATURES

PREFACE

RESEARCH TECHNIQUES

The Laboratory and its Techniques

Divisions of Research

PROSODIC FEATURES

1. Declarative Intonation
2. Non-Declarative Intonation
3. Place of Logical Stress in the Word
4. Place of Logical Stress in the Sense Group
5. Nature of Logical Stress
6. Place of Emphatic Stress
7. Nature of Emphatic Stress
8. Variations in Syllable Weight
9. Internal Juncture and Syllabication
10. Syllable Types
11. Tension

VOCALIC FEATURES

1. Articulatory Description
2. Acoustic Description
3. New Vowel Sounds for the English Speaker
4. Distribution (Positional and Allophonic)
5. Frequency of Occurrence
6. Duration System
7. Neutral Position
8. Loss of Color
9. Effect of Consonant Anticipation on Vowels
10. Diphthongization
11. Effect of Syllable Type on Vowel Color
12. Attack and Release
13. Nasality

CONSONANTAL FEATURES

1. Articulatory Description
2. Acoustic Description
3. New Consonant Sounds for the English Speaker
4. Distribution (Positional and Allophonic)
5. Frequency of Occurrence
6. Duration System
7. Neutral Position
8. Consonantal Weakening
9. Effect of Vowel Anticipation on Consonants
10. Speed of Articulation
11. Tongue Fronting
12. Aspiration
13. Affrication
14. Palatalization
15. Final Release
16. Voicing

PREFACE

In order to improve the teaching of modern languages in our schools, we need to know much more about the general phonetic characteristics that distinguish English from the languages that are to be taught. These general phonetic characteristics are not just the individual vowels and consonants of a language, or the phonemic systems they form, but rather those broad acoustic or articulatory features that are common to all or most sounds or sound sequences of a language, such as: tenseness or laxness of articulation, closed or open syllabication, diphthongization of vowels, aspiration of consonants, affrication of plosives, fronting or backing of the tongue, rounding or spreading of the lips, sharp or weak vowel onset, released or unreleased consonant offset, loss of vowel quality under unstress, nasalization, syllabic type, syllabic equality, syllabic rate, rhythmic patterns, stress patterns, intonation patterns, duration patterns, the nature of logical stress, the nature of emotive stress, etc.; in a word, those phonetic features that are the key to good pronunciation, to an effective use of the phonemic system of a language.

This volume summarizes the present state of our knowledge after five years of research under Contracts SAE-8366 and OE-2-14-026 with the United States Office of Education.

In a first chapter, the four-way technique of research we have developed for the objective comparison of languages is described, together with the electronic instrumentation we have constructed. Then, in the three following chapters, the prosodic features, the vocalic features and the consonantal features of English, French, German and Spanish are compared under forty subdivisions (see Table of Contents).

During the past two years, under Contract OE-2-14-026, major advances have been made in the following fields, as reflected by the list of publications and papers found at the end of this Preface.

Our film library of X-ray motion pictures of tongue profile has been enriched for French and English.

We have made articulatory comparisons of vowels and consonants by X-ray profile views of the tongue and front views of the lips. Films of lip movements for /e/, /o/, /i/, /u/ vowels in comparative sentences in English, French, German and Spanish have been assembled, as well as X-ray films of the tongue profile for the two types of American /r/, called retroflexed and bunched.

A detailed comparison of diphthongization in English and French was made by spectrographic analysis and synthesis, and by motion pictures of lips correlated with X-ray profiles of tongue-motion sequences for /i/, /e/, /o/, /u/. This study was published in the French Review.

As a sequel to the two articles previously published on Spanish and French declarative intonation as compared with English, a detailed comparison of German and English declarative intonation was continued and will soon be finished and prepared for publication.

The comparison of syllable types, of phonemic frequency of occurrence, of syllable length as a factor of rhythmical patterns, have progressed considerably.

The factors of consonant voicing have been isolated by synthesis for English.

New correlates of the consonant/vowel distinction have been discovered by means of an electronic gate.

A phonetic comparison of German with French and English has been made which shows some fourteen features in common between German and French, and only six between German and English.

In addition, some advances have been made in each of the forty features of languages that are being investigated.

A list of articles published in professional journals and of papers read at professional meetings between June 15, 1962, and June 15, 1964, under contract OE-2-14-026 with the Office of Education follows. (Other articles are being prepared for publication or are in course of publication.)

Articles

"Voyelles diphtonguées et voyelles pures." The French Review 37, 1: 64-76, October 1963.

"Research techniques for the phonetic comparison of languages." IRAL, 1,2: 85-97, 1963.

"Comparing the prosodic features of English, German, Spanish and French." IRAL, 1/3 and 4: 193-210, 1964.

"Quality in tape recording and voicing." International Journal of American Linguistics, 29,2: 55-60, 1963.

Papers

"Pitch levels versus pitch shapes in teaching English, German, Spanish and French intonation." Rocky Mountain Modern Language Association Meeting, Utah, 1962.

"A cross-linguistic comparison of syllables." RMMLA Meeting, Utah, 1962.

"A comparison of the frequency of phoneme occurrence." RMMLA Meeting, Utah, 1962.

"English phonetics as heard by speakers of German, French, and Spanish." Meeting of Western Conference of Foreign Student Advisers, Boulder, 1963.

"The complexity of voicing." Meeting of the American Association of Teachers of Speech, Denver, 1963.

"Distinctive and non-distinctive aspects of German intonation." RMMLA Meeting, Denver, 1963.

"Cineradiographic evidence of retroflex and bunched varieties of American r." RMMLA Meeting, Denver, 1963.

"Interference of American diphthongization in teaching the pure vowels of French, German and Spanish." RMMLA Meeting, Denver, 1963.

"Change as a correlate of the consonant/vowel distinction." Linguistic Society of America Meeting, Chicago, 1963.

"German phonetics between French and English." MLA Meeting, Section of General Phonetics, Chicago, 1963.

"Experimental films for the phonetic comparison of languages." University of Kentucky Foreign Language Meeting, April 1964.

RESEARCH TECHNIQUES FOR
PHONETIC COMPARISON OF LANGUAGES¹

INTRODUCTION

The teaching of modern languages in our schools and colleges would certainly improve if we could describe very objectively the new phonetic habits that are to be acquired by a student when, or before, he learns a second language. To reach this goal, advantage must be taken of the extraordinary technical advances made during the last decade or two in speech research centers, mainly the Haskins Laboratories of New York, but also the Massachusetts Institute of Technology Acoustics Laboratories of Boston, the Bell Telephone Laboratories of Murray Hill and the Royal Institute of Technology Speech Transmission Laboratory of Stockholm. While the phonemicists were absorbed in the formulation of rigorous methods for the structural analysis of languages as systems, some experimental phoneticians achieved gigantic progress, in those laboratories, concerning the acoustic cues -- components of distinctive features -- that correlate with the perception of vowels, consonants, and prosody. First, they analyzed the acoustic elements of speech by measurement on spectrograms. Spectrography makes the complex patterns of syllables accessible to the naked eye in a three-dimensional display of time, frequency and intensity.

¹ The research reported herein was performed pursuant to a contract with the U.S. Office of Education, Department of Health, Education, and Welfare.

and shows a different pattern for every vowel, every consonant, as well as dialectal and individual traits of pronunciation or marks of a foreign accent. On spectrograms, differences of duration, intensity, pitch, syllabication, diphthongization, affrication, aspiration, nasalization, vowel color, consonantal place of articulation, and so on, can be located and measured. Figure 1 will perhaps give an idea of what can be observed on a spectrogram to those who are not familiar with this kind of speech display. Second, experimental phoneticians synthesized the acoustic features of speech. All the information found on spectrograms is not linguistically relevant; some of it, for instance, refers to voice quality. It became necessary, therefore, to construct machines that could play back the spectrographic patterns of speech so that one could make omissions and modifications and judge the effects by ear. Such artificial speech synthesizers turned out to be ideal tools for research. By transforming hand-painted patterns into very intelligible speech they made it possible to isolate the acoustic cues of speech perception, to vary their dimensions of duration, frequency, and intensity, and determine their optimum values. This research technique has been so productive that a trained spectrographer can today paint, free hand, patterns of speech that will produce whatever dialectal traits he wishes, when passed through the synthesizer. Finally, experimental phoneticians were curious to relate the constantly changing patterns portrayed on spectrograms to the constantly changing

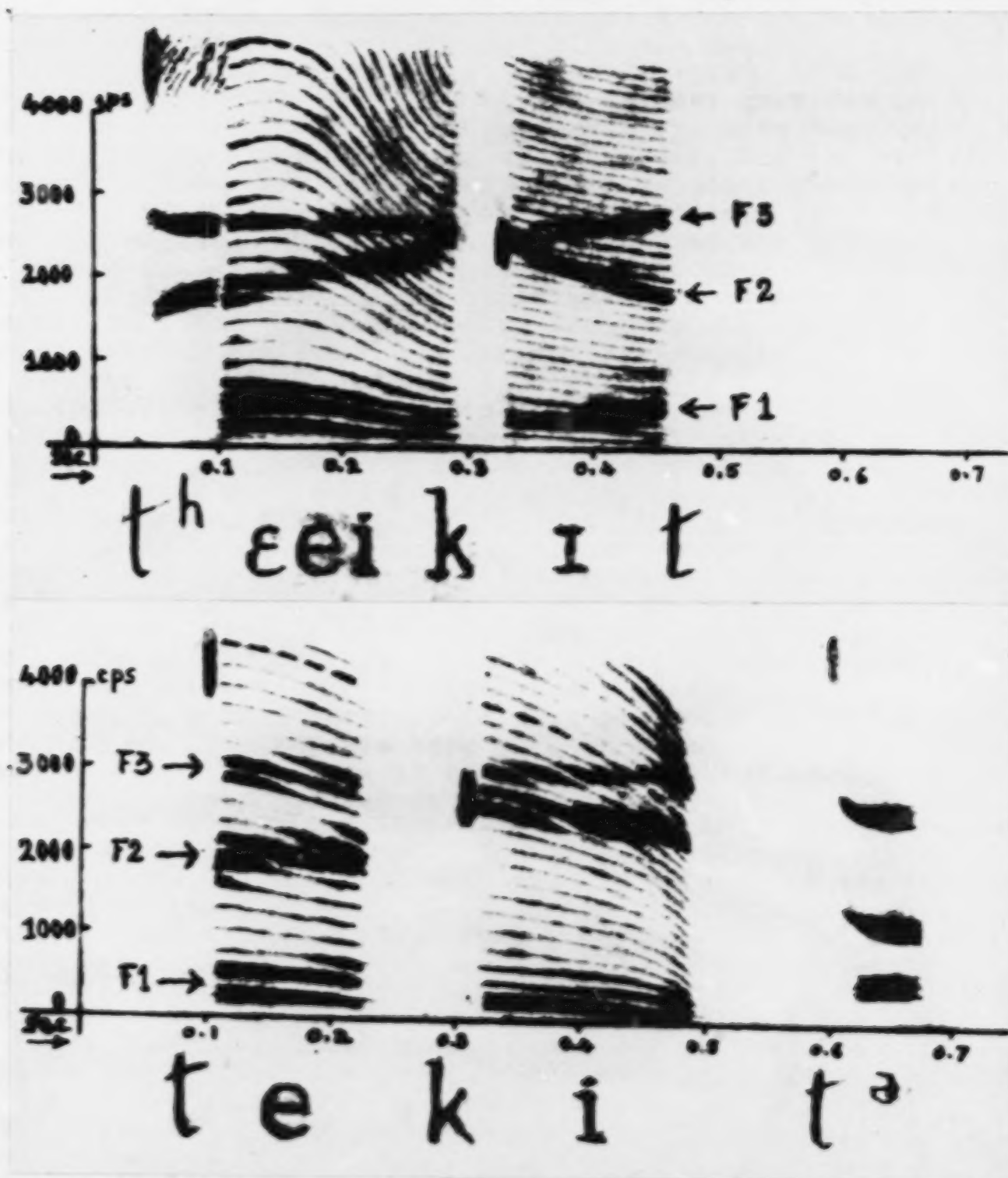


Figure 1

(Caption for Figure 1)

Schematized spectrograms of the command "Take it" are shown as uttered by an American (above) and by a Frenchman (below). The thin lines represent the rich range of harmonics produced by the vocal cords. Those harmonics are thicker (more intense) whenever their frequency coincides with the continuously changing notes of resonance of the vocal tract cavities. From left to right, the darker lines formed by the thicker portions of the harmonics are the "formants." Phonemically, note that the first formant (F1) is higher, and the second lower, for [e] or [I] than for [i]; for [t]'s, the bursts are relatively high and the second and third formants originate respectively (or point to) lower and higher than the center frequency of the [e]'s; for [k]'s, the bursts are at the second-formant onset, therefore lower than for [t]'s, and the second and third formants tend to meet. In general, only the first and second formants are necessary to recognize vowels, but the third-formant transitions play an important part in the identification of consonants. Phonetically, we note the following differences between the two speakers. 1. Place of stress. The American stresses the first syllable; the Frenchman the last. 2. Nature of stress. The American shows more syllabic duration, more syllabic intensity, more aspiration, higher pitch (the dashed line) on the first syllable; the Frenchman shows more syllabic duration on the last syllable, but not more intensity, or higher pitch. 3. Intonation. It can be followed from left to right on any harmonic, but the higher the harmonic the easier the reading of pitch. The high harmonic that is marked with dashed lines indicates that in order to express command the American uses a concave, reversed-S shape; the Frenchman uses a convex shape. Both shapes are falling and could be called 2-14 phonemically, yet they are very different phonetically. 4. Aspiration. The American shows it before the stressed vowel; the Frenchman shows it nowhere. Before the unstressed vowel, the American [k] is lenis -- although there are no vocal cord vibrations (the first harmonic is not present) the closure is short enough to create a semi-impression of voicing. 5. Release of the final consonant is shown in French; the American speaker does not show any in this rather typical case. Note that voiceless [e] of consonant release has a first formant, whereas [h] of aspiration does not (this is a phonemic, not a phonetic remark). 6. American diphthongization of [e] is recognizable by the falling frequency of the first formant and the rising frequency of the second formant [eɪ] in their middle sections (the faster rises at the beginning and end of the second formant belong not to the diphthongization of the vowel but to the perception of the [t] and [k]). 7. Consonant anticipation is visible in [tɛɪk] by the joining of the arresting transitions of the second and third formants by the American; the arresting transitions of the French [k] join less than its releasing transitions after the [k] burst.

completed, except for such refinements as research alone can

indicate, and considerable research has already been accomplished. cavities of the vocal tract, and they developed the use of motion picture X-rays, to replace the still X-rays. To reduce drastically the dangers of radiation, light intensifiers were invented. Then, to verify theories of speech production based on cineradiography, artificial mouths were built -- and electric analogs of the vocal tract -- which produced excellent syllables by changing only three parameters: the mouth constriction, lip constriction and cavity volumes. As a result, the basic theory of speech production has become clear and simple. It involves a source of sound and cavity resonance of that sound. In vowels the source is at the glottis; for consonants it may be at any point of constriction between the glottis and the lips.

It occurred to us that the same technique as, for instance, that which was used at Haskins Laboratories to study the acoustic correlates of speech perception could be applied to the study of foreign accents, that is, the phonetic features that really make the difference between one language and another. Thus the most advanced electronic techniques would serve a practical purpose: better teaching of foreign languages. Fortunately, instrumental techniques of artificial speech synthesis were as familiar to us as the problems of language teaching. Under contract with the U. S. Office of Education, we undertook in 1959 the construction of appropriate equipment for the phonetic comparison of languages. By now this construction is basically completed, except for such refinements as research alone can

lower formants of [x] to their frequency position for [a]).

indicate, and considerable research has already been accomplished.

raises the frequency of the friction, which is higher in French because the [THE LABORATORY AND ITS TECHNIQUES apical]. Such

Our research equipment for speech can be described under three headings: acoustic analysis, acoustic synthesis, and articulatory study by motion-picture.

For acoustic analysis, we adapt sound spectrographs to our particular purpose. To the Kay Sonographs we add a frequency marker, by means of a complex tone oscillator, to assure perfect calibration for all conditions of heating or zero line; a cooling system; and external vernier controls for the zero line. We modify the time scales of 5 in. per second to make it correspond to standard tape recording and to our synthesizer scale of 7.5 in. per second. To the 2000 cps per inch frequency scale we add one at 1200 cps to correspond to our synthesizers, and another at 200 cps, which enlarges the lower harmonics by ten times, to facilitate the study of intonation curves (the undulating rise and fall of the first or second harmonic generally offers a direct reading of frequency variations). We also connect to the sonograph two accessories: one is a variable-frequency power supply for speed-control of the sonograph turntable motor. It permits judging by ear and measuring objectively the frequency changes that result from varying the playback speed of a speech segment. One can observe the conditions under which an English [ɛ] changes to a German [ø] (a slight slowing down lowers both lower formants of [ɛ] to their frequency position for [ø]),

or a Spanish [s] changes to a French [s] (a slight speeding up raises the frequency of the friction, which is higher in French because the [s] sounds are more fronted and less apical). Such speed-control would also show frequency relations among fricatives: as the frequency of a [ʃ] noise is raised by continuous acceleration, it is heard successively as [s], [θ], and [f].

The other accessory is a sound-gating device. This device makes it possible to separate from an utterance any portion as short as one centisecond or as long as 2400 centiseconds, and to listen to it without erasing the original utterance. Thus one can hear separately the various parts of a diphthong. One can also find the optimum point of separation between vowel and consonant and determine to what extent the consonant-to-vowel transition belongs to the perception of the consonant. This is done by displacing at will the point of separation and listening to both portions separately.

The preceding device operates on time-segments (vertical sections on a spectrogram). It is also possible to separate frequency-bands from an utterance (horizontal layers on a spectrogram). For this we use the analyzer part of the vocoder that we have built ourselves. This analyzer divides the utterance into twenty frequency bands, any one of which can be omitted in order to judge its contribution to the total perception. Thus it is found, for instance, that practically all vowel information is below 3000 cps, but the relevant consonant noises of the [s], [θ], [f] type are well above that.

It should be noted here that separating formants from an

utterance is very different from separating frequency bands. Formant frequencies rise and fall, and cut across the bands of frequency. In order to isolate whole formants and judge their contribution to speech perception, one needs not only an analyzer but an artificial speech synthesizer -- a far more involved piece of equipment.

For acoustic synthesis we have constructed two sizable machines, both of which permit the isolation and manipulation of the acoustic factors of speech, the acoustic correlates of distinctive features as well as, for instance, the acoustic elements of a foreign accent. These two machines are conceived on different electronic principles and complement one another. Synthesizer I has 50 channels, the 50 harmonics of a series, produced by a tone wheel. Because of the great number of channels it is efficient in the use of formant transitions and produces very sharp consonants, but it cannot vary its fundamental frequency and speaks in a monotone. Synthesizer II uses the synthesizer part of a Vocoder and has only 18 frequency channels. Its consonants are not so sharp but it can vary the fundamental by simple hand-painting of frequency curves, making possible the experimental study of pitch variations (intonation). Both machines play back spectrographic patterns of speech by light reflection. The patterns can be made on spectrographs or be painted by hand. They must be white to permit reflection of the light from a powerful arc.

These two synthesizers are at the heart of our research equipment. Observing a difference between a German and a

Spanish syllable on a spectrogram permits no more than hypothesizing. In order to ascertain the relevancy of this difference as a cross-linguistic contrasting factor, it is necessary to synthesize the utterance with and without the feature in question and compare the results by ear. To a considerable extent, then, the synthesizers permit the verification of every hypothetical difference between two languages, in isolation or in combination. It eliminates much of the speculative aspect that has marked this kind of research in the past.

The third section of our laboratory is for motion-picture study of the tongue, jaws, velum action taken by X-ray, or of lip action taken normally. In order to observe the movements of articulation that take place inside and outside of the mouth, as well as the positions of articulation, on motion-picture films, we have two types of viewers: 1) projectors with sound, which must be used at regular speed (24 frames per second) if the sound is to be heard without distortion; to facilitate observation, we have added stop, go and reverse push-buttons to such projectors; 2) projectors without sound but with a hand crank permitting one either to stop at each frame or to watch at will in forward-backward motion short sequences of frames. On X-rays, considerable detail is made visible by motion which cannot be detected on a stopped frame. It is a combination of stop and motion which reveals the most. With two hand-cranked projectors, one can compare on the same screen the articulatory positions of a similar syllable in two languages, frame by frame or sequence by sequence. In addition to sound and silent projectors, we

have film readers that project the frames on a flat surface, where tracing the X-ray shadows or the lip contours becomes possible. We are also able to print frames in any reasonable size.

Motion-picture X-rays are not taken in our laboratories but in appropriate hospitals that are equipped especially for that purpose and know what dosage of exposure a subject can receive without danger of radiation. X-ray laboratories use the language materials (lists of words and sentences) that we have prepared for comparison of English to foreign languages point by point. For each language these materials total about two pages and take three to four minutes to be spoken, a time that is too long for a single person. Here are examples of utterances for the comparison of English [e] and [o] to Spanish [e] and [o] in stressed position: It's a day still to face, Que le dé dos cafés; It's LOW pay to cope with, Don Lope da la copa. Conditions of stress and phonic environment are kept as similar as possible in the two languages.

To summarize, our research technique appears to be threefold: spectrographic analysis, spectrographic synthesis, and motion-picture analysis. This, however, is only from the instrumental aspect. A fourth category, which might be called statistical, plays an important role in showing how certain features prevail quantitatively. Statistical study applies, for instance, to vowels (What are the most frequently occurring vowels? How frequently do we meet those vowels that are very characteristic of a given language and not found in others?), consonants, clusters, syllable sounds, syllable types (open, close), syllable length, place of stress, falling intonations, etc.

DIVISIONS OF RESEARCH

Preliminary investigations have led us to divide our comparison of English to Spanish, German and French into forty areas of phonetic difference. This does not mean exactly that so far we have found forty phonetic differences between English and any one of these three target languages -- although this is not completely unrealistic -- but that we have found it convenient to do our investigating separately under that many headings. Time will tell how many further subdivisions will be necessary. This dividing system is of interest in itself -- we shall present the greatest part of it here with brief indications of the research techniques applicable to each area. Results that have already been obtained will be presented in later articles. Let it be clear that we are interested in all phonetic features of English and the target languages that can help in teaching the latter.

1. GENERAL COMPARISON OF VOWEL COLOR. To compare the vowel phonemes of a language to those of another in a manner that will be useful to the teacher, both the articulatory and the acoustic description must be presented. Articulatory. a) The profile of the tongue and lip positions for each vowel must be obtained not from X-ray stills, as in the past, but from X-ray motion pictures which show the same position with much more constancy. After comparing several subjects and taking into consideration the influence of various contiguous consonants, schematized tongue and lip shapes must be drawn. b) Phonetic charts of tongue position

position but on lip shape and any other articulator responsible for the volume of vocal tract cavities. A comparison of acoustic charts among languages will give a true picture of relative qualities of vowel systems: their richness, their balance, their center of gravity, their regions of concentration such as low, front, back, center, etc. It will show relations, for instance, between duration and place of production, nasality and place of production. This is explained by the fact that lip rounding enlarges and closes the front cavity thereby lowering the frequency of the second formant, and that rounded vowels make their constriction slightly less fronted than the corresponding spread vowels, which reduces the volume of the back cavity and raises the frequency of the first formant. Acoustic. Vowels may be considered as musical chords, mainly characterized by the frequency (pitch) of their two lowest formants. Spectrograms show those two formants (different for every vowel) and make it possible to measure their frequency. a) The spectrographic aspect of the whole vowel system of each language can be schematized. b) An acoustic chart of vowels, quite similar to the triangular or quadrilateral articulatory chart based on the highest point of the tongue, can be made for each language by graphing on logarithmic paper the first formant frequencies in ordinate and the second formant frequencies in abscissa. Such acoustic charts have many advantages over the articulatory charts. They are more objective since formant frequencies are more subject to measurement than somewhat unstable tongue shapes; they represent more complete information since they rely not only on tongue

position but on lip shape and any other articulator responsible for the volume of vocal tract cavities. A comparison of acoustic charts among languages will give a true picture of relative qualities of vowel systems: their richness, their balance, their center of gravity, their regions of concentration such as low, front, back, center, etc. It will show relations, for instance, between duration and place of production, nasality and place of production. (by decreasing the back cavity volume). But those

[A technical parenthesis must be made here, however. The frequencies of characteristic formants cannot be obtained only by measurement of spectrograms; they must be verified by synthesis. To arrive at truly representative figures for a standard voice pitch (fundamental frequency) artificial vowels must be tested by ear. This is necessary because formant frequencies vary somewhat with individual voices. The formants of a feminine voice can be 10 to 15 percent higher than those of a male voice. Besides, verification by synthesis will prevent errors of all sorts that have crept into too many spectrographic studies of late, errors of formant location, errors of measurement, and errors due to faulty adjustment of the spectrograph itself.]

Correlation. The acoustic and the articulatory are naturally related. a) Formant frequencies actually are the notes of resonance of the vocal tract, which is generally divided into two cavities, one in the pharynx (back cavity) and one in the mouth (front cavity). The larger its volume, and the smaller and longer its opening, the lower the note of resonance of a cavity. Roughly speaking, therefore, observing changes in the

frequencies of first and second formants is like observing modifications in volume and openings of the back and front cavities, respectively. b) Some practical relations may be more useful to the teacher than the one above: tongue fronting raises the second formant (by decreasing the front cavity volume); lip rounding lowers the second formant (by increasing the volume and closure of the front cavity); and tongue lowering raises the first formant (by decreasing the back cavity volume). But those relations are indirect. For instance, whenever there is tongue lowering, in series such as [i, e, ε, æ], [y, ø, œ, a], or [u, o, ɔ, ɑ], as the tongue lowers, the place of constriction moves toward the back until it is far down into the pharynx for the [ɑ] family, leaving only a small back cavity between mid-pharynx and the glottis. It is the backing of the place of constriction which directly decreases the volume of the back cavity, thereby raising its resonant frequency. The lowering of the tongue is only an indirect factor. c) One can see now that the relation between the articulatory chart based on the highest point of the tongue and the acoustic chart based on the frequencies of formants is an indirect one: backing and lowering the highest point of the tongue changes the place of constriction, which in turn changes the volume and opening of the resonating cavities.

Dealing somewhat at length with a resonance aspect of vowel production will help us understand more rapidly many other features of speech.

2. GENERAL COMPARISON OF CONSONANTS. The perception of consonant noises, like that of vowels, depends on cavity resonance

in the vocal tract. But whereas the source of sound (the sound to be resonated) is always far back at the glottis (vocal cords) for vowels, it can be anywhere between the glottis and the lips for consonants, and only the portion of the vocal tract that is located in front of the consonant constriction resonates the consonant noise quite efficiently. Consonant noises are mainly distinguished by their frequency: [ʃ] noise is lower than [s] noise, but higher than [x] noise; and that frequency depends on the volume and closure of the cavity in front of the constriction. In a general way, therefore, the farther back the constriction, the larger the resonance cavity and the lower the consonant noise.

The perception of consonants, however, does not depend only on consonant noise. It relies also on the rapidly changing formant frequencies (formant transitions) which reflect the rapid movements of articulation taking place from vowel to consonant or consonant to vowel. Consonant is movement; vowel is steady-state. In a syllable [do] for instance, only the relatively steady-state end of the formants belong to the vowel; the rapidly changing formant transitions that can be seen on a spectrogram between the [d] explosion and the [o] steady-state belong entirely to the perception of the [d]. Articulatory. To compare consonants among languages, a teacher will first need a) complete sets of profiles of tongue and lip positions as well as b) a schematized profile view of all the places of articulation for each language. X-ray motion pictures will provide this information, but not X-ray stills. Acoustic. a) A teacher also needs a chart of the frequencies of consonant noises in two

parts: one for friction noises such as in [f, θ, s, ʃ, ç, x, X], another for burst noises such as in [p, t, c, k]. Correlation of these noise frequencies with the place of articulation profile will enable the teacher to show, for instance, why the [s] sounds of English are lower (in frequency) than those of French, and those of Spanish lower yet. b) Acoustic charts showing the role of formant transitions in the perception of consonants are probably too involved to have any practical value in teaching.

The frequencies of noises will be suggested by spectrographic analysis, but, as in the case of vowels, they will require confirmation by synthesis in artificial syllables.

3. NEW SOUNDS. A language teacher will welcome a separate list, with acoustic and articulatory descriptions, of the sounds that are radically different from English in the target language, as well as a list of the sounds that are radically different from the target language in English. Acoustic and articulatory charts of these new sounds will bring out the regions and types of articulation that are the most characteristic and cause the greatest interference in learning.

4. DISTRIBUTION. With the phonemes of a language, a teacher needs to know a) in which positions they occur, b) how they vary according to position. The former will show, for instance, that short vowels and voiced plosives are not found in final position in German; the latter that postvocalic [l] is not retroflex in French.

5. SOUND FREQUENCY. To complement the information conveyed by comparative charts of distinctive sounds and their distribution,

it is indispensable to know also the frequency of occurrence of each sound. a) The acoustic impression made by the target language on a new student depends on which sounds recur more than others, what the proportion of vowels to consonants is. Initial teaching material should perhaps make use of such knowledge.

b) The importance given to "new sounds" in the early stages of learning should be considered in the light of their frequency of occurrence. Frequency counts must be based on conversational or dramatic material as well as on narrative material, and if the two types are found not to diverge markedly the results can be combined.

6. SYLLABLE CHARACTERIZATION. Another frequency count that may throw some light on the phonic characterization of a target language concerns a) the syllables that recur most often, their actual phonemic composition, their proportion of vowels to consonants; b) types that prevail in terms of "open syllables" (CV, CCV) or "closed syllables" (CVC, VC); c) the length of syllables, in terms of the number of phonemes, and the relation of length to frequency of occurrence and to speed of articulation. These last two areas are strictly statistical.

7. TYPICAL TONGUE SHAPE. Each language seems to have a characteristic shape of the tongue, either in key consonants and vowels or in rest positions to which the tongue tends to return. Motion picture X-rays may make it possible to compare such a factor among languages.

8. THE NEUTRAL VOWEL. When an American child hesitates, the vowel he utters is mid-central to mid-open-back with lips nearly

unrounded. It is close to the unstressed schwa of sofa. A French child uses a mid-front well rounded vowel under the same circumstances. X-ray views of such articulatory positions should be very useful in teaching. This neutral vowel is a sort of home base to which the tongue returns frequently in the course of speech.

9. TENSION. Objective measurement of the degree of tension or laxness in the muscles of articulation has not yet been possible. If such a thing exists as a language characteristic, it might correlate with such factors as the proportion of steady-state vs movement, the speed of movement, the time it takes to pass from fast change to relative steady-state, etc., which can be measured on spectrograms and tested by ear in synthesis.

10. DURATION SYSTEMS. Variations of vowel duration are of two very different kinds: those that are conditioned, automatic, unconscious and nondistinctive, such as the effect of stress, the rate of utterance, the influence of the following consonant -- [I] is shorter in mitt than in mid; and those that are learned and contribute to a phonemic distinction, such as in German Stadt vs Staat, in English bit vs beat, in French tache vs tâche. A teacher must be aware of both kinds in order not to confuse what should be taught and what should not. Measurements of length are made on spectrograms and verification of length perception by synthetic speech. The same two types of length variations are found in consonants: in intervocalic position single flap vs multiple flap is distinctive in Spanish, [karo/karro], whereas it is not in Italian [ieri-ierri].

11. DECLARATIVE INTONATION SHAPES. These include the two basic and most frequent modes of expression: continuation and finality. Analysis. To compare them among the four languages, we use spontaneous speech by intelligent speakers. Hundreds of feet of spectrograms are made in the 2000 cps per inch scale to show formants and high harmonics and in the 200 cps per inch scale to permit one to follow the first or second harmonic, which offers a direct reading of frequency variations. From a statistical study of such undulations, we extract the most typical shapes. Synthesis. Then we paint those shapes on Synthesizer II and judge by ear the effect of certain changes and omissions to discover which parts of the shapes are most relevant and what is the subjective impression (pitch) of those objective shapes (frequency) in each language. We also impose the pitch shapes of one language on the words of another language to judge the degree of characterization. We would like our pitch shapes to be the phonetically relevant schematization of the objective frequency shapes. They are not to be confused with the subjective pitch levels (2-3-1↓) which can be distinctive within a language but tend to mask differences among languages. Phonetic shapes, which vary sharply from language to language, may be more useful than phonemic levels which use somewhat the same digital combinations in all languages.

12. NONDECLARATIVE INTONATION SHAPES. These will include frequent attitudes such as command, question, interrogation, implication, parenthesis, echo, exclamation. The techniques of analysis are the same as for declaration. but to be markedly different in languages with variable place of stress.

13. THE PLACE OF LOGICAL STRESS. Here, the method applied is statistical. We analyze a few thousand running words of narrative and conversational material in each language, and we extract for the teacher two sorts of information: a) how variable or predictable the place of word stress is; b) the characteristic place of stress in one-, two- three-, and four-syllable words. In a later chapter on rhythmic patterns, we will apply nearly the same method to finding the place of stress in sense groups.
14. THE NATURE OF LOGICAL STRESS. To discover the relative contributions of excesses of intensity, duration, and frequency (objective factors) to giving prominence (subjective impression) to a syllable, the speech synthesizers are incomparable tools. The three objective factors can be varied individually and in combination, and the effects judged by ear by natives of each language. Spectrographic analysis is limited, here, to making hypotheses.
15. THE PLACE OF EMPHATIC STRESS. In many languages it is different from the place of logical stress. In French, for instance, emotive stress bears on the first or second syllable, whereas logical stress is on the last. Statistical methods are applied.
16. THE NATURE OF EMPHATIC STRESS. Pitch and intensity are used more here than in logical stress. The teacher needs to know how differently these ingredients are mixed in the target language and in his native language. The synthesizers are the main research tools.
17. RHYTHMICAL PATTERNS. A statistical study of the place of stress in sense groups is also necessary, because word rhythm (Chapter 14) and sense-group rhythm turn out to be markedly different in languages with variable place of stress.

18. PHONETIC SYLLABICATION. It must be studied at two levels:

a) its phonetic realization when it is distinctive -- the German contrast zum einen vs zu meinen is made quite differently from the English an aim vs a name; b) its phonetic characteristics when it is nondistinctive -- English aiding vs French aider.

A synthesizer makes it possible to isolate the acoustic correlates of syllabication and to vary them at will. They contribute considerably to making a phonic sequence sound German, English, Spanish, or French.

19. VARIATIONS IN SYLLABLE WEIGHT. Here we study the distribution of intensity, duration and frequency in all syllables of a text, unstressed as well as stressed, in relation to such factors as their position in the phrase, or their grammatical function. This involves spectrographic measurements, statistics, and verification by synthesizers of the character it contributes to each language. In English, for instance, final syllables in unstressed position are found to have less intensity but as much or more duration than non-final stressed syllables.

20. LOSS OF VOWEL COLOR. To what extent, in each language, do vowels retain the same color under different degrees of stress? How different are the o's in abolish and abolition? Here vowel formants are being measured on spectrograms.

21. DIPHTHONGIZATION. Languages vary considerably from one another in the way so-called monophthongs actually change in color. Two factors seem relevant here: the amount of change, and the direction of change (toward center, toward ends). These changes can profitably be transferred from spectrograms to vowel

mentally by spectrographic analysis, vocal cord study, and charts. They require a thorough acquaintance with consonant spectrography to avoid confusing vowel changes with formant transitions that are related to the perception of consonants.

22. DIPHTHONGS. Spanish and German have three diphthongs in common with English: [ai], [au], and [oi]. But the phonetic realization of these diphthongs is different in each language in at least three ways: the color of initial and target vowels, the relative amount of steady-state of each portion, and the decrease of intensity from beginning to end. Spectrograms, correlated with X-ray motion pictures of the tongue movements plus front- and side-view motion pictures of the lip movements, are being studied.

23. VOWEL NASALITY. Its phonetic realization as a distinctive feature (as in French) must be studied separately from its occurrence as a nondistinctive characterization (as in American English or in Spanish). Both can be seen on spectrograms as well as on X-rays.

24. VOWEL COLOR AND SYLLABLE TYPE. The type of syllable -- closed or open -- does not affect a vowel to the same extent in all languages. For instance, the tendency for ai to open in closed syllable (j'aime) and to close in open syllable (j'ai) is very marked in French. Similar tendencies also exist in other languages, but to a lesser extent. Spectrographic measurements of formants will specify the differences.

25. VOWEL ATTACK AND RELEASE. The sharp attack, with or without vocal cord explosion, is used more in Germanic than in Latin languages. Various degrees of attack have to be studied experimentally with actual plosives and actual fricatives in the other lan-

mentally by spectrographic analysis, vocal cord study, and vowel synthesis.

26. TIME DIMENSIONS OF CONSONANTS. One character of the consonants that differs from language to language is relative speed of closing, duration of closure, and speed of opening. These dimensions are being studied by motion pictures of lip and tongue (X-ray) movements in comparative sentences such as: Two galley bees are dozing in a candy pail, vs Tout son cadeau est un habit payé à la gare. To make use of the maximal extent of formant transitions, velars must be before central vowels, dentals before back vowels, and labials before front vowels, as in these sentences.

27. CONSONANT FRONTING. The degree of tongue-tip fronting for all the apical consonants contributes markedly to the characteristic overall resonance of a language, for it affects not only the consonants themselves but the contiguous vowels as well. It is being studied by means of X-rays and spectrograms in comparative sentences such as: Take the same native shale from the lake in the dale, vs Les aînés ont aidé chez les Leger l'été passé.

28. CONSONANT ASPIRATION. How late the glottis remains open after the plosive burst of noise can be measured on spectrograms and reproduced on synthesizers. Comparative sentences such as The publication of a text... vs La publication d'un texte... are used.

29. CONSONANT FRICATIVATION. This is mostly found in Spanish, when voiced plosives stand between vowels. The acoustic and articulatory nature of those fricativated plosives needs to be compared with actual plosives and actual fricatives in the other lan-

guages, also between vowels. Comparative sentences: Dan was good about bidding long ago, vs Goberno el abogado sin duda en Vegas.

30. CONSONANT AFFRICATION. Affrication can be distinctive as in tico vs chico in Spanish, or nondistinctive as in English tick, kit vs French tic, quite with less friction after the burst than in English. To study this phenomenon we correlate X-rays with spectrograms.

31. CONSONANT PALATALIZATION. English tends to combine a yod with a contiguous lingual consonant: miss you is often pronounced [mɪʃu]. Other languages may behave quite differently under the same circumstances. X-rays and spectrograms make the comparison possible among languages.

32. CONSONANT RELEASE. The degree of release of final consonants differs from language to language. Spectrograms show this and indicate how consonant closure correlates with it. Tests of the contribution of release to perception of place of articulation can be made by synthesis as well as by tape-splicing manipulation.

33. CONSONANT ANTICIPATION. Post-vocalic consonants show a tendency to close earlier in some languages than in others, and after certain vowels than after others: German bitte vs biere. Arresting transitions of the formants show this by their length and their rate of change. It can also be studied in motion pictures. Marked consonant anticipation gives a jerky impression to the foreign listener.

34. VOWEL ANTICIPATION. The degree to which one takes the position of a vowel while, or even before, articulating the

preceding consonant, varies considerably among languages.

Sentences such as the following permit motion picture comparison between English and German: The boat sailed down the Po,
vs Das Boot schwimmt auf dem Po.

35. CONSONANT VOICING. It appears that initial [d] voices earlier in Spanish than in English and conversely intervocalic [t] voices more in English than in Spanish. Such voicing and unvoicing conditions in all positions are being compared among languages as they pose important problems of interference in learning. The acoustic and articulatory correlates of voicing are found to be extremely complex, however, and cannot be limited to the presence of a "voice bar" on spectrograms.

SUMMARY AND CONCLUSION

Our comparative study of prosody is in progress under the following: 1. Our purpose is the improvement of language teaching. We have described our four-way research technique -- spectrographic analysis, spectrographic synthesis, articulatory motion-picture study, and statistical calculation -- for the comparison of the phonetic characteristics of English, Spanish, German, and French; and we have very briefly introduced the application of such techniques to some thirty-five phonetic differences among those languages -- all differences of considerable importance in the teaching of those languages. 2. In later articles the actual results that have already been obtained will be presented in three parts: vowels, consonants, and prosody.

1. The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

COMPARING THE PROSODIC FEATURES OF ENGLISH, GERMAN, SPANISH AND FRENCH¹

This is the second in a series of articles on the general phonetic characteristics of American English as compared with those of German, Spanish and French -- the three languages that are taught most extensively in America. It is assumed here that the American teacher of German, Spanish or French who is to benefit from this research is familiar with the phonemic systems of those languages but needs to know more about their phonetic features in order to teach more effectively.

The first article described at length our technique of research. It was presented as four-fold -- spectrographic analysis, spectrographic synthesis, radiographic observation, and statistical survey. This one is devoted to prosody, or non-segmental features. Later ones will compare the segmental features of the four languages. As our research is still in progress, present results can only be partial, but long range views are also discussed.

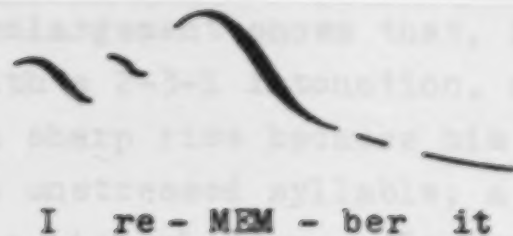
Our comparative study of prosody is in progress under the following subdivisions: 1. Declarative intonations. 2. Non-declarative intonations. 3. The place of logical stress in the word. 4. The place of logical stress in the sense group. 5. The nature of logical stress. 6. The place of emphatic stress. 7. The nature of emphatic stress. 8. Variations in syllable weight. 9. Juncture and syllabication. 10. Syllable types. 11. Tension.

I. DECLARATIVE INTONATION

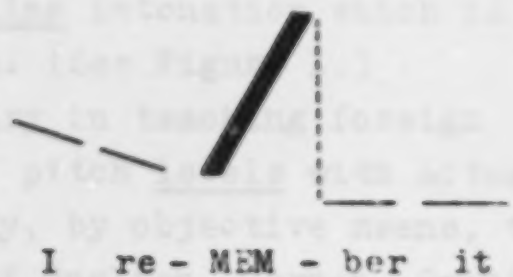
The use of pitch levels, or pitch phonemes, often indicated by digits (Es regnet: 2-3-1; It's raining: 2-3-1) has been found quite insufficient in teaching the intonation of a second language. Only a native of the language can interpret the digits correctly. The learner of a second language needs

1. The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

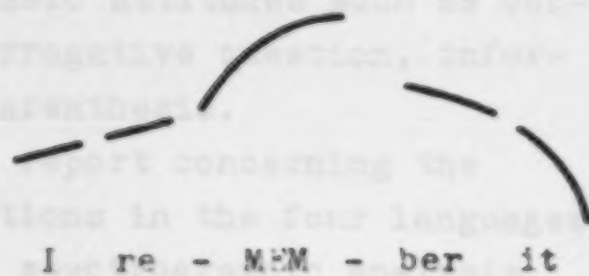
American intonation 2-3-1



Spanish imitation
of American 2-3-1



French imitation
of American 2-3-1



German imitation
of American 2-3-1



Figure 1

much more information than mere levels if he is not to impose his own pitch shapes on each syllable. For instance, our pitch research by spectrographic enlargement shows that, in attempting to say I remember it with a 2-3-1 intonation, a Spaniard will convert the 3 into a sharp rise because his habit is to fall only after a high unstressed syllable; a Frenchman into a rise followed by a high plateau; a German into an S-shaped rise starting with a slight dip. None of them will produce the "reversed-S" falling intonation which is so characteristic of American English. (See Figure 1.)

Realizing that it was necessary in teaching foreign languages to complement the use of pitch levels with actual shapes, we have undertaken to study, by objective means, the most characteristic pitch shapes of English, German, Spanish and French for the expression of basic attitudes such as continuation, finality, command, interrogative question, informative question, implication and parenthesis.

At present we have results to report concerning the pitch shapes of declarative intonations in the four languages under study. They were obtained by spectrographic analysis of at least five minutes of spontaneous speech by cultivated natives of each country.

To obtain a well defined pitch line, we amplify by a factor of ten the lowest frequency portion of each spectrogram (see Figure 2). This enlarges the frequency variations of the low harmonics (which are generally continuous for voiced sounds) and makes their observation relatively easy. We usually examine the rise and fall of pitch by following the first or second harmonic, whichever one shows more clearly at a given moment. (We assume in this study that the subjective impression of "pitch" variations nearly correlates with the objective measurement of "frequency" variations. Cases in which length and intensity factors may interfere with this correlation have not yet been clearly determined by psycho-acousticians.) Then we make a statistical study of the various shapes that occur and recur for the expression of each

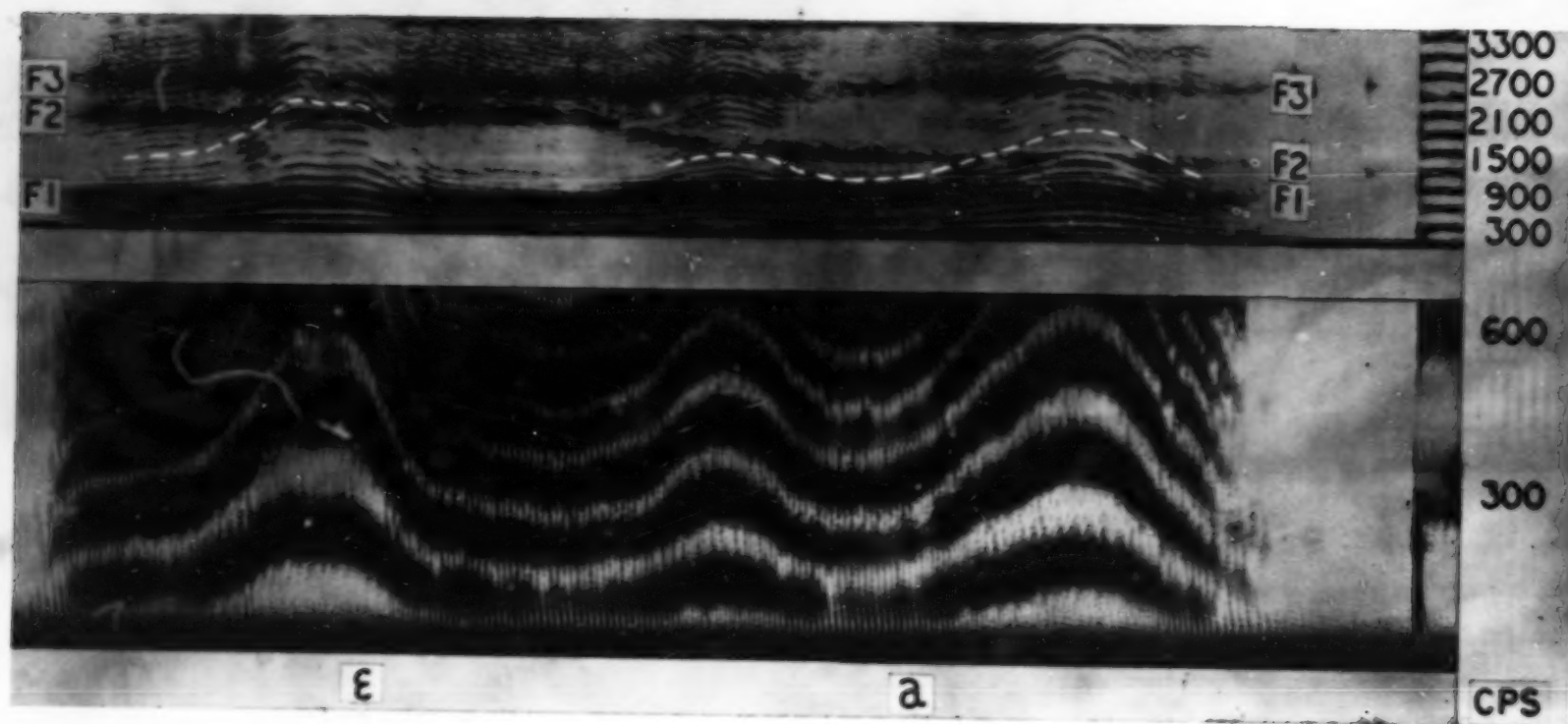


Figure 2

attitude. Finally, we synthesize, on a speech synthesizer of the voback type, the prevailing shapes that have emerged, and we determine what portions, what traits of those shapes, are essential to their characterization, by making omissions or changes and judging the effect by ear. We can also superimpose the pitch shapes of one language on the words of another, or even test the intonation shapes of the four languages by using an identical nonsense utterance for all of them.

Figure 3 presents the prevailing pitch shapes we have found for minor continuation, major continuation, and finality in English, German, Spanish and French. In addition to frequency variations of the fundamental, intensity and duration variations are also present. French is quite different from the three other languages in that it shows no variations in intensity, and has the stress (by greater duration) on the ultima. The range of frequency, indicated by the fixed horizontal lines on the figure, is of 7 to 10 tones, depending on the subject rather than on the language.

As we compare pitch shapes among the languages in Figure 3, let us consider each of the four sentences as being divided into three sense groups. On the whole, English clearly stands apart from the three other languages; but marked differences are also apparent among the three. Let us observe one feature at a time.

1. In contrast to the other languages where rising portions prevail, falling portions are predominant in the American utterance, not only in the stressed syllables but in the unstressed ones as well. It is this constant recurrence of falling glides that characterizes American intonation for the ear of foreigners. Furthermore, the falling portions of stressed syllables are the ones that bear the greatest intensity (loudness); the rising portions, whether leading to the slope or following it, are weak and not very perceptible. In synthesized intonation, the pre-slope rise can be omitted without reducing the American characterization.

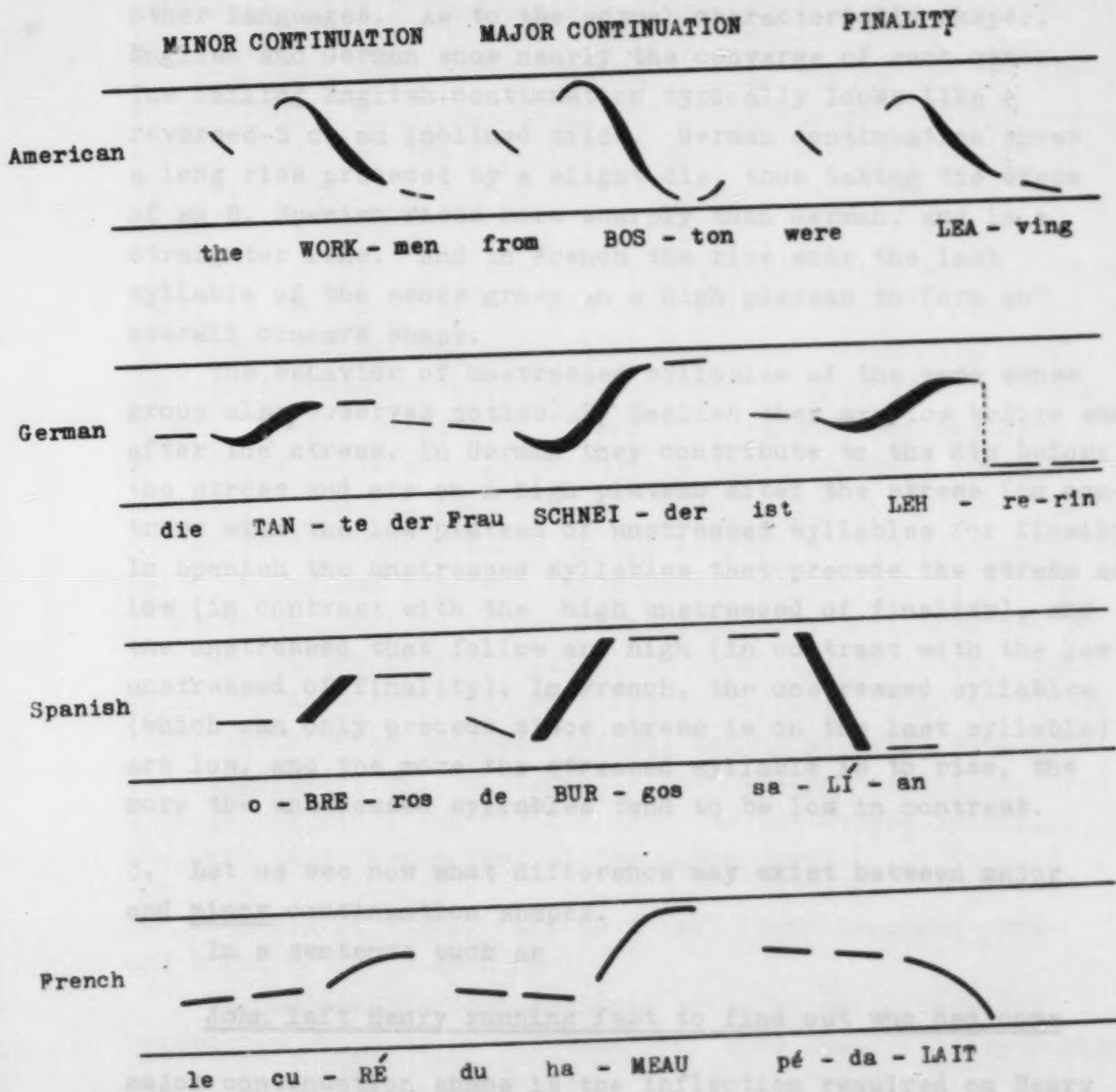


Figure 3

2. Continuation (major as well as minor) is mainly expressed by falling shapes in English, and by rising shapes in the three other languages. As to the actual characteristic shapes, English and German show nearly the converse of each other. The falling English continuation typically looks like a reversed-S or an inclined tilde. German continuation shows a long rise preceded by a slight dip, thus taking the shape of an S. Spanish rises more sharply than German, and in a straighter line. And in French the rise ends the last syllable of the sense group on a high plateau to form an overall concave shape.

The behavior of unstressed syllables of the same sense group also deserves notice. In English they are low before and after the stress. In German they contribute to the dip before the stress and are on a high plateau after the stress (in contrast with the low plateau of unstressed syllables for finality). In Spanish the unstressed syllables that precede the stress are low (in contrast with the high unstressed of finality), and the unstressed that follow are high (in contrast with the low unstressed of finality). In French, the unstressed syllables (which can only precede since stress is on the last syllable) are low, and the more the stressed syllable is to rise, the more the unstressed syllables tend to be low in contrast.

3. Let us see now what difference may exist between major and minor continuation shapes.

In a sentence such as

John left Henry running fast to find out who had come

major continuation shape is the inflection required on Henry to help make clear that John is running, not Henry (in which case a minor continuation shape covers fast):

John left Henry // running fast / to find out who had come ‡

or on fast to help indicate that Henry is running, not John

(in which case a minor continuation shape covers Henry):

John left Henry / running fast // to find out who had come #

In this case, major continuation on fast unites the two minor sense groups, John left Henry and running fast. (Naturally, pause and stress can also play a part in identifying the one who runs, but we are only concerned here with the role of pitch variations.)

The contrast between minor continuation and major continuation is much more marked and much more frequently observed in the three target languages than in English; only very good English speakers observe it consistently. That is not to say that it should be ignored in English, as phoneticians have been doing -- we have found subtle evidence of it in many examples of spoken American. As shown in Boston (Figure 3), the major continuation has on the average a more pronounced fall (longer and steeper) and a larger proportion of rising hooks at the end. (A larger proportion of rising, instead of falling, curves is also found.) In German, Spanish and French, major continuation rises more frequently than minor continuation; the rise is steeper, higher, and/or preceded by a lower dip (Schneider, Burgos, hameau). The high end of the stressed rise is sustained and carried onto the unstressed syllables (if any) that form a part of the sense group.

4. Finality is mainly falling in all four languages. But how differently in each one!

In English the last stressed syllable takes the same reversed-S shape as in continuation, with unstressed syllables (if any) prolonging the slope.

In German, after the S-shaped glide, there is a sharp break of the unstressed portion (generally a series of syllables) to a flat low level which gives the impression of being separated from what precedes. A similar effect of separation would be produced in English if, instead of street-car level, we heard street // car-level.

In Spanish, on the other hand, finality is announced well ahead of the stressed syllable by a high unstressed portion -- the pre-stress syllables of the final sense group. The stressed syllable falls steeply from that high level to a very low one for the unstressed syllables that may follow. In spite of the resemblance to German, it is quite different because in Spanish the fall is on the stressed syllable itself and the unstressed syllables which follow remain closely connected to it.

In French the pattern of finality begins on the first syllable of the last minimal sense group, not on the last stressed syllable, as in the three other languages; and it begins lower than the preceding continuation rise had ended unless there is "accent d'insistance."

Different from the three other languages, the shape of the finality curve is convex in French. This is in direct contrast with the typically concave ending shape of English finality.

5. The contrast between continuation and finality is not so clearly realized in English as in the other languages. Both continuation and finality have falling shapes; the difference is often merely a matter of degree in the slope or of frequency range. This explains why foreigners constantly mistake an American continuation for a finality. They find it unambiguous only when the rising hook is clearly voiced, and this occurs much less than the majority of the time.

In the three other languages the contrast "continuation/finality" is quite unambiguous: continuation rises, finality falls. This is clearest in Spanish: not only is the fall on the last stressed syllable, but the preceding unstressed syllable unmistakably announces that fall by a high pitch. In French there is no ambiguity either, but nothing warns you clearly ahead of time; only the last syllable (always stressed) contains the definite slope of finality. In German the stressed syllable of the final sense group begins with a rising S-shape somewhat similar to that of continuation; it is only the following unstressed syllables that clearly indicate finality.

II. NON-DECLARATIVE INTONATION

In comparing among our four languages the typical pitch shapes for such attitudes as command, interrogative question (rising), information question (falling), implication, and parenthesis, we use the same electronic techniques of spectrographic analysis and synthesis, but the material is quite different. Extemporaneous speeches by good speakers do not frequently contain such intonations. We have to use, on one hand, extemporaneous discussions recorded from symposia, interviews, or from radio programs; and on the other hand, recorded conversations in which we create schemes to elicit responses with varied intonations from native subjects. No results can be reported at this time.

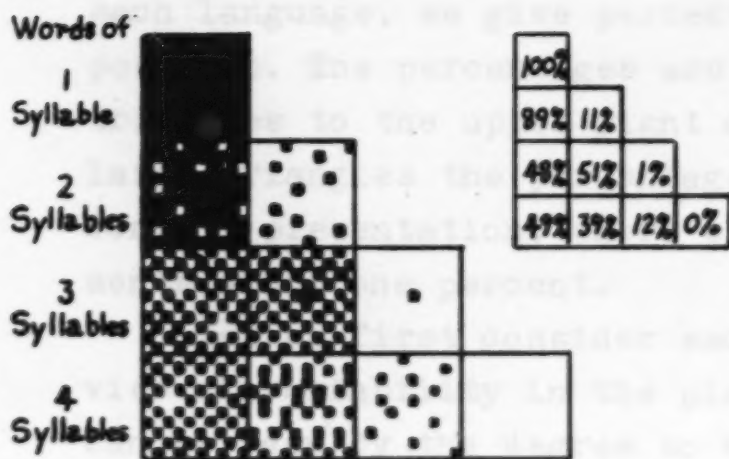
III. THE PLACE OF LOGICAL STRESS IN THE WORD

In teaching a language, it is necessary to know: a) the characteristic place of stress; b) how variable or predictable this place is. This is related to stress patterns in words and in sense groups. For practical purposes of cross-linguistic comparison, we must first limit our investigation to primary stress in words.

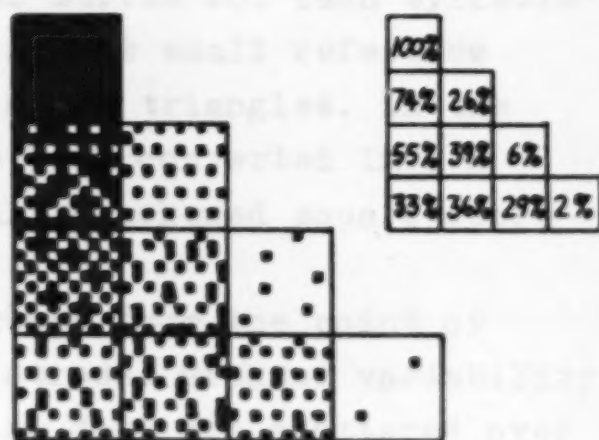
A preliminary study of the position of primary word stress was made through an analysis of contemporary short story texts. Short stories in French, Spanish and German were used, with a translation of these three texts as an example of English.

Our statistics are based on running words from texts of five to eight pages in length, and not vocabulary count. There were approximately 1500 words analyzed for both French and Spanish and 2400 for German. The larger number of English words (5800) is due to the fact that all three translations representing the other language texts were used.

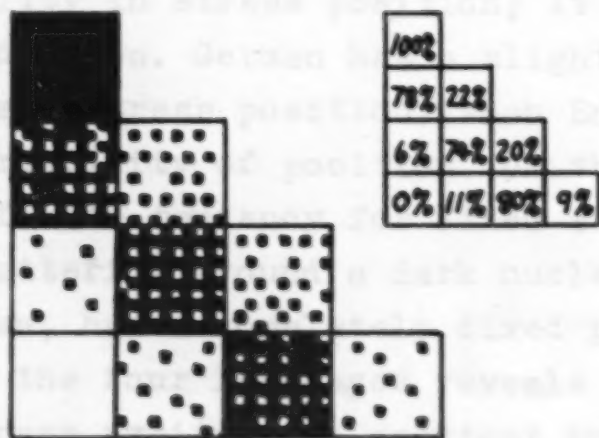
Figure 4 presents the results of our study in schematic form. Because of the small number of polysyllabic words of five or more syllables in our count, we have limited our comparison to words of one through four syllables. In the figure,



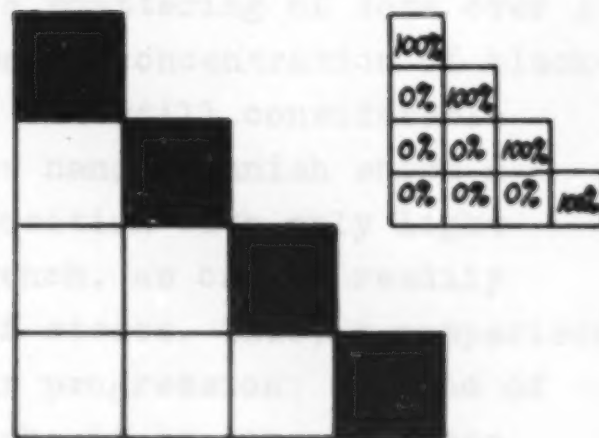
German



English



Spanish



French

German
English
Spanish
French

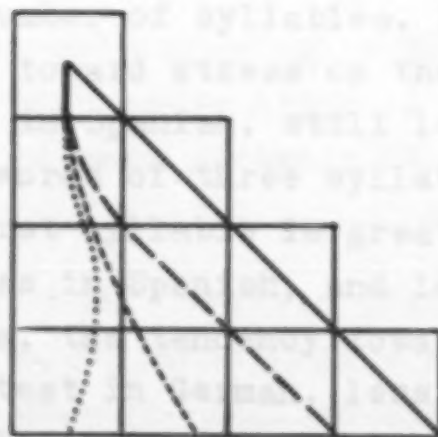


Figure 4

syllable position within words reads from left to right. For each language, we give percentages of stress for each syllable position. The percentages are given in the small reference triangles to the upper right of the major triangles. In the large triangles the percentages have been converted into a density presentation, where each small blackened square represents about one percent.

Let us first consider each language from the point of view of variability in the place of stress. Greater variability can be seen by the degree to which the dots are scattered over the triangle, while more predictable stress is shown through a concentration of dots in a specific area. With this in mind, we see that, of the four languages, English has the greatest variability in stress position; it shows a scattering of dots over a wide area. German has a slightly greater concentration of blackness (stress position) than English, but still considerable variability of position. On the other hand, Spanish shows a definite tendency for fixed stress position with only light scattering around a dark nucleus. French, as can be readily seen, has a completely fixed place of stress. Thus, a comparison of the four languages reveals a clear progression: freedom of stress position is greatest in English; it is somewhat less marked in German; a tendency toward a fixed position appears in Spanish; and a completely fixed position is reached in French.

Let us now compare the place of stress in the four languages. It varies according to the number of syllables. In words of two syllables, the tendency toward stress on the first syllable is greatest in German, less in Spanish, still less in English, and least in French. In words of three syllables, the tendency toward stress in the first syllable is greatest in English, less in German, still less in Spanish, and least in French. In words of four syllables, the tendency toward stress in the first syllable is greatest in German, less in English, less again in Spanish, and least in French. The tendency in four-syllable words coincides with the general tendency as summarized in the lower right-hand corner of Figure 4. On

the whole, German tends to place the stress on the first syllable. English does also, but to a lesser degree. Spanish strongly tends to have it on the penult, and French has it without exception on the ultima.

The correlation between place of stress and variability is not complete:

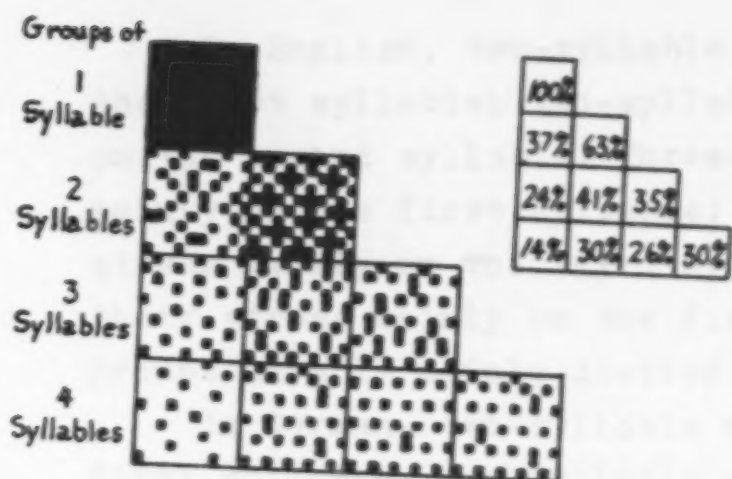
Tendency toward placing the stress near the beginning decreases in the order: German, English, Spanish, French. Tendency toward varying the place of stress decreases in the order: English, German, Spanish, French.

IV. THE PLACE OF LOGICAL STRESS IN THE SENSE GROUP

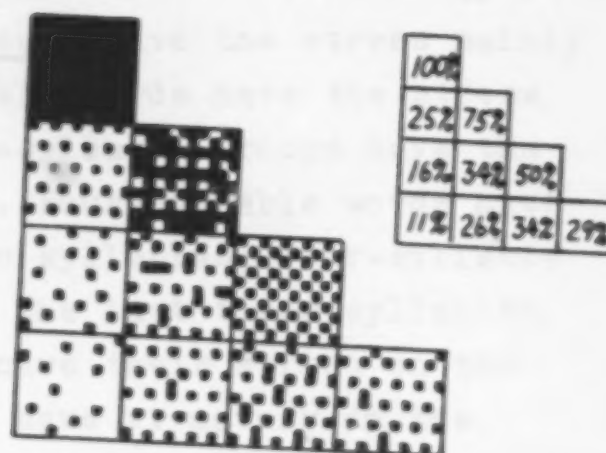
Here we present a preliminary comparison among the four languages of the place of primary stress (only) in short sense groups. A further comparison between the word patterns of Part III and the sense group patterns of this part will be instructive.

The same three short story texts and their English translations were used for our analysis of sense groups as for our study of word stress. They were divided into sense groups as nearly as possible on the same principle for each of the four languages. The statistical results of primary stress patterns are presented in schematic form in Figure 5. The sense groups had to be very short, for no practical patterns could emerge from longer divisions.

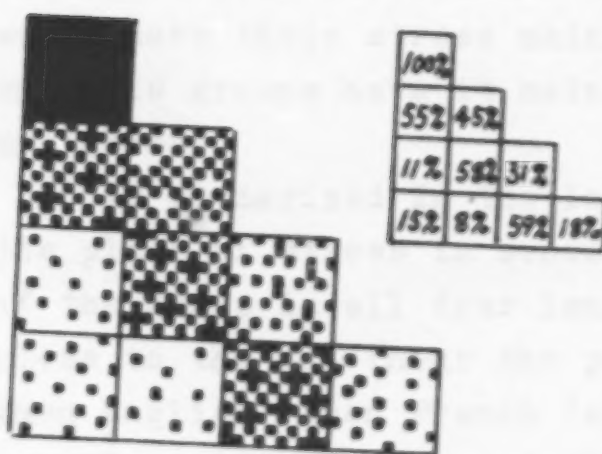
1. The scattering of dots is considerably greater in Figure 5 than it was in Figure 4. This means that sense group patterns of rhythm do not recur as regularly as word patterns of rhythm. It could have been expected.
2. The second result was quite unexpected. Whereas Spanish stress remains on the penult syllable, in sense groups as in words, although in a less conclusive manner, English and German stresses are severely drawn toward the end of the sense group -- and slightly more so in English than in German.



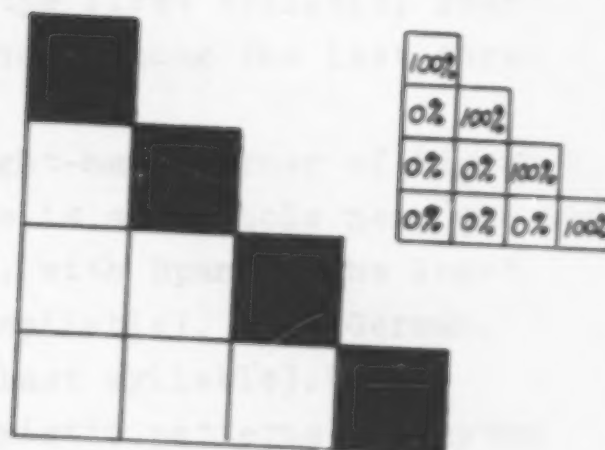
German



English



Spanish



French

German
English
Spanish
French

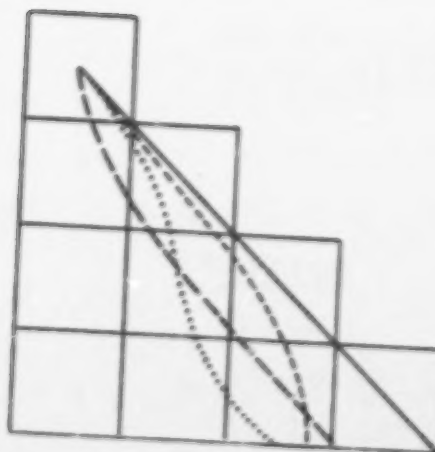


Figure 5

In English, two-syllable words have the stress mainly on the first syllable; two-syllable groups have the stress mainly on the second syllable. Three-syllable words have the stress mainly on the first syllable; three-syllable groups have the stress mainly on the third syllable. Four-syllable words have their stress mainly on the first two syllables; four-syllable groups have it mainly divided among the last three syllables.

In German, two-syllable words have their stress on the first syllable; two-syllable groups have it mainly on the second syllable. Three-syllable words have their stress divided between the first two syllables; three-syllable groups have it mainly divided between the last two syllables. Four-syllable words have their stress mainly on the first syllable; four-syllable groups have it mainly divided among the last three syllables.

As summarized in the lower right-hand corner of Figure 5, the place of stress in sense groups is as a whole near the end of the group in all four languages, with Spanish the least close to the end (near the penult syllable), then German, then English, then French (on the last syllable).

Examples of the most characteristic patterns of rhythm for each length of words and each length of sense groups follow.

2 syllables	<u>lickers</u>	<u>with light</u>
3 syllables	<u>possible</u>	of the <u>past</u>
4 syllables	<u>inhabited</u>	at the <u>corner</u>
	<u>yellowish</u>	<u>distressing scenes</u>

<u>French</u>		
1 syllable	<u>choses</u>	<u>certain</u>
2 syllables	<u>tristesse</u>	en <u>moi</u>
3 syllables	<u>impression</u>	et la <u>parole</u>
4 syllables	<u>débarrasser</u>	de la <u>nature</u>

V. THE NATURE OF LOGICAL STRESS

	<u>Word</u>	<u>Sense Group</u>
	<u>Spanish</u>	
1 syllable	<u>sé</u>	<u>que</u>
2 syllables	<u>fueron</u>	<u>juntos</u> <u>la tez</u>
3 syllables	<u>momentos</u>	<u>en casa</u> <u>al volver</u>
4 syllables	<u>inocentes</u>	<u>a la puerta</u> <u>des corazón</u> <u>fué preciso</u>
	<u>German</u>	
1 syllable	<u>Milch</u>	<u>Ach</u>
2 syllables	<u>Finger</u>	<u>er sass</u>
3 syllables	<u>heftiger</u> <u>betrogen</u>	<u>beseelt war</u> <u>wie der Wind</u>
4 syllables	<u>eingeschlummert</u>	<u>gestanden war</u> <u>zu erwarten</u> <u>auf seinen Wunsch</u>
	<u>English</u>	
1 syllable	<u>give</u>	<u>first</u>
2 syllables	<u>flowers</u>	<u>with light</u>
3 syllables	<u>possible</u>	<u>of the past</u>
4 syllables	<u>inhabited</u> <u>melancholy</u>	<u>at the corner</u> <u>distressing scenes</u>
	<u>French</u>	
1 syllable	<u>choses</u>	<u>certes</u>
2 syllables	<u>tristesse</u>	<u>en moi</u>
3 syllables	<u>impressions</u>	<u>et la perte</u>
4 syllables	<u>débarrasser</u>	<u>de la nature</u>

Preliminary experiments on complete sentences (synthesized) indicate that in English, German, and Spanish, the main factor

V. THE NATURE OF LOGICAL STRESS

Stress being what makes a syllable stand out, our experiments have concentrated on separating the acoustic elements that cause a syllable to be linguistically perceived as prominent. At this point we find that syllabic prominence is related to variations in:

- a) vowel intensity
- b) vowel duration
- c) consonant duration (length of closure -- it correlates with a subjective impression of consonant intensity)
- d) consonant intensity (energy in the noise of burst, as in [p] sound, or the noise of friction, as in [s] sound)
- e) syllabic pitch (fundamental frequency variation)

In synthesis, any of these five factors can be studied in isolation, and any combination of the five can be produced. Research by playback and voback is very promising, but we have only exploratory results to report here.

When we synthesize an utterance like nice party and reverse the ratios of a, b, c, e, one at a time in the two syllables of party, that is, when we successively give the syllable ty the vowel intensity, the vowel duration, the consonant duration, or the syllabic pitch of par, and conversely, the impression of stress unambiguously shifts from par to ty. However, it does not sound equally like "American stress" in each case.

Such experiments can be repeated with words that change meaning when the stress is shifted, such as suspect (noun) which becomes suspect (verb). The advantage of such words is that they can be judged phonemically by naive listeners who, instead of answering the question "which syllable is stressed," answer the question "is it a noun or a verb." Such words, unfortunately, never have comparable consonants -- belonging to the same class -- such as /p/ and /t/ in the word party.

Preliminary experiments on complete sentences (synthesized) indicate that in English, German, and Spanish, the main factor

of syllabic prominence is a combination of vowel intensity and consonant closure duration, whereas in French it is vowel duration. We also have indications that vowel duration plays a greater role in German stress and a smaller role in Spanish stress than in English stress. And pitch is important in all four, perhaps more in English than in the three target languages.

Our aim is to discover in what proportion the five ingredients a, b, c, d, e, must be mixed to produce a characteristic stress in each of the four languages. This will involve a great number of experiments by synthesis, with judgement tests by native speakers of each language.

The nature of stress has been one of the most investigated problems of phonetics. But all previous investigations have failed to include the factor of consonants which is being taken into account here. Pitch changes may supersede increases of vowel intensity, but they do not supersede so easily a combined increase of consonant duration-intensity and vowel duration-intensity as they occur in speech.

VI. THE PLACE OF EMPHATIC STRESS

Special emphasis on a word tends to shift the stress toward the beginning of the word if it is not there already. This is a regular process in French where impossibilité in logical stress becomes IMpossibilité or imPOssibilité in emphatic stress, depending on how emotive the reason for emphasis is. The tendency also exists in English, German, and Spanish where the stress is occasionally shifted to an earlier syllable for emphasis on the word. We are comparing the extent of this phenomenon in the four languages. Habitually, however, English, German, and Spanish, unlike French, keep emphatic stress on the same syllable as logical stress.

a) open syllables: 20 centiseconds; b) closed syllables (more consonants): 25 cc.

VII. THE NATURE OF EMPHATIC STRESS

Here, the ratios between a) vowel intensity, b) vowel duration, c) consonant duration, d) consonant intensity, and e) syllabic pitch, will not be the same as for logical stress. The role of vowel duration is expected to be reduced and that of consonant duration to be increased. Pitch variations may also play a more important role. We hope to find how these factors will vary from language to language.

VIII. VARIATIONS IN SYLLABLE WEIGHT

Syllable weight, like stress, is a combined function of loudness, pitch, and length, which can be measured by intensity (decibels), frequency (cycles), and duration (centiseconds). We have already spoken of the nature of stress. Here we wish, if possible, to go a step further and compare among the four languages not only what makes the stressed syllables stand out, but also how the unstressed syllables behave in regard to each other and in regard to the stressed ones; in regard to their position in the phrase and in regard to their grammatical function. This will involve measuring variations of intensity, frequency, and duration of all the syllables along the phrase. In the end it should yield comparative schematic representations of prosodic features that will show at a glance the all-important differences in the distribution of intensity, frequency, and duration along the sentence.

Our first exploration is limited to variations of syllable length in English. We have measured the syllables of five minutes of extemporaneous speech by Margaret Mead -- over 3000 syllables in all.

Very roughly, five main categories of syllable duration emerge. Results are given here in round numbers.

1. Stressed syllables that are not final in a phrase,
 - a) open syllables: 20 centiseconds; b) closed syllables (more consonants): 25 cs.

2. Stressed syllables that are final in a phrase,
a) open syllables: 40 cs; b) closed syllables: 45 cs.
3. Unstressed syllables of monosyllables that are not final in a phrase (mainly words like the, we, of, than),
a) open syllables: 6 cs; b) closed syllables: 10 cs.
4. Unstressed syllables of polysyllables that are not final in a phrase, a) open syllables: 11 cs; b) closed syllables: 15 cs.
5. Unstressed syllables that are final in a phrase,
a) open syllables: 25 cs; b) closed syllables: 30 cs.

This is shown in Figure 6.

It is interesting to note that final unstressed syllables are on the average equal to or perhaps even longer than non-final stressed ones. One must note, however, that the long final unstressed syllables have comparatively low vowel intensity and short consonant closure to compensate for their long duration. Therefore, long syllabic duration does not exclusively correlate with heavy syllabic weight in English, as strong vocalic intensity and long consonant closure regularly do.

IX. INTERNAL JUNCTURE AND SYLLABICATION

Languages differ in their junctural habits, especially in regard to internal juncture (for instance, in connected word separation: it swings / its wings, a nice man / an ice man) and syllabication (syllable separation within a word: compare French inabilité to English inability; every syllable of the French word is more open, shows less consonant anticipation, than the corresponding syllable of the English word). Here we shall limit our discussion to internal juncture in the consonant-vowel or vowel-consonant sequences. In English, this type of transition can be distinctive (it can function so as to produce a change of meaning) by its presence or absence: an aim [ən+eɪm], a name [əneɪm]; an ice man [ən+aɪsmən], a nice man [ənaɪsmən]; or by its change of place: great eye

VARIATIONS OF SYLLABLE LENGTH
IN ENGLISH

In German, internal syllable boundaries can also be distinctive: *zu weinen* is different from *zu einen*, *so leben* from *woll leben*.

But in French and Spanish, internal syllable boundaries are not distinctive (at least in the normal spoken language). The syllable is not different from the word. The syllable is the unit of pronunciation. The syllable is the unit of the phonetic system. The syllable is the unit of the phonetic system. The syllable is the unit of the phonetic system.

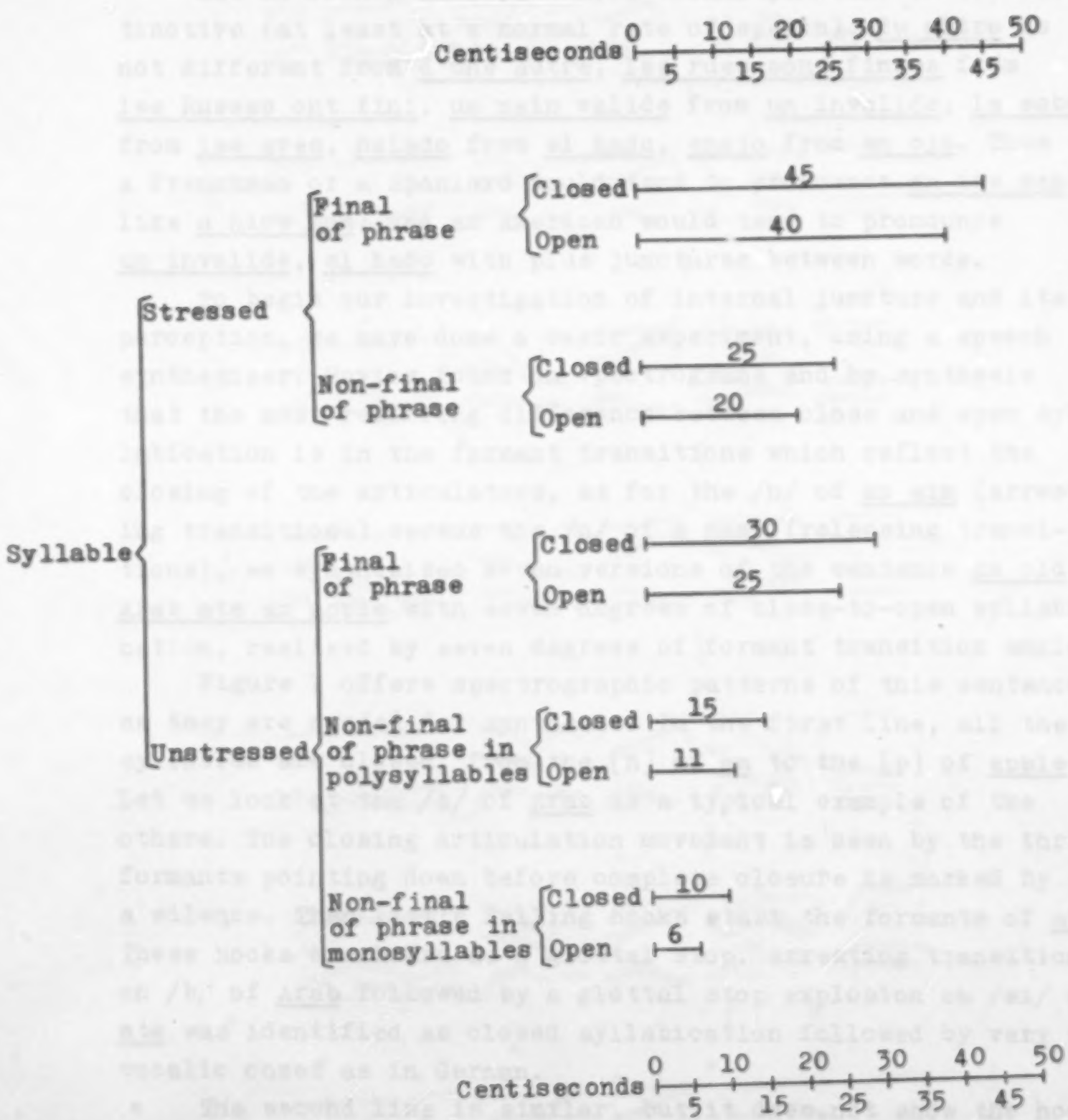


Figure 6

[greit+ai], gray tie [grei+tai]; house ad [haus+əd], how sad [hau+səd].

In German, internal juncture can also be distinctive: zu meinen is different from zum einen, wo leben from wohl eben.

But in French and Spanish, internal juncture is not distinctive (at least at a normal rate of speech): du nôtre is not different from d'une autre, les rues sont finies from les Russes ont fini, un nain valide from un invalide; la sabes from las aves, helado from el hado, enojo from en ojo. Thus a Frenchman or a Spaniard would tend to pronounce an ice man like a nice man; and an American would tend to pronounce un invalide, el hado with plus junctures between words.

To begin our investigation of internal juncture and its perception, we have done a basic experiment, using a speech synthesizer. Having found on spectrograms and by synthesis that the most recurring difference between close and open syllabication is in the formant transitions which reflect the closing of the articulators, as for the /n/ of an aim (arresting transitions) versus the /n/ of a name (releasing transitions), we synthesized seven versions of the sentence An old Arab ate an apple with seven degrees of close-to-open syllabication, realized by seven degrees of formant transition angles.

Figure 7 offers spectrographic patterns of this sentence as they are needed for synthesis. In the first line, all the syllables are closed, from the [n] of an to the [p] of apple. Let us look at the /b/ of Arab as a typical example of the others. The closing articulation movement is seen by the three formants pointing down before complete closure is marked by a silence. Then little falling hooks start the formants of ate. These hooks are heard as a glottal stop. Arresting transitions on /b/ of Arab followed by a glottal stop explosion on /ei/ of ate was identified as closed syllabication followed by very sharp vocalic onset as in German.

The second line is similar, but it does not show the hooks of glottal stops at the beginning of ate. The vocalic attack is less sharp. This was identified as closed syllabication followed by fairly sharp onset as in English.

an old Arab ate an apple

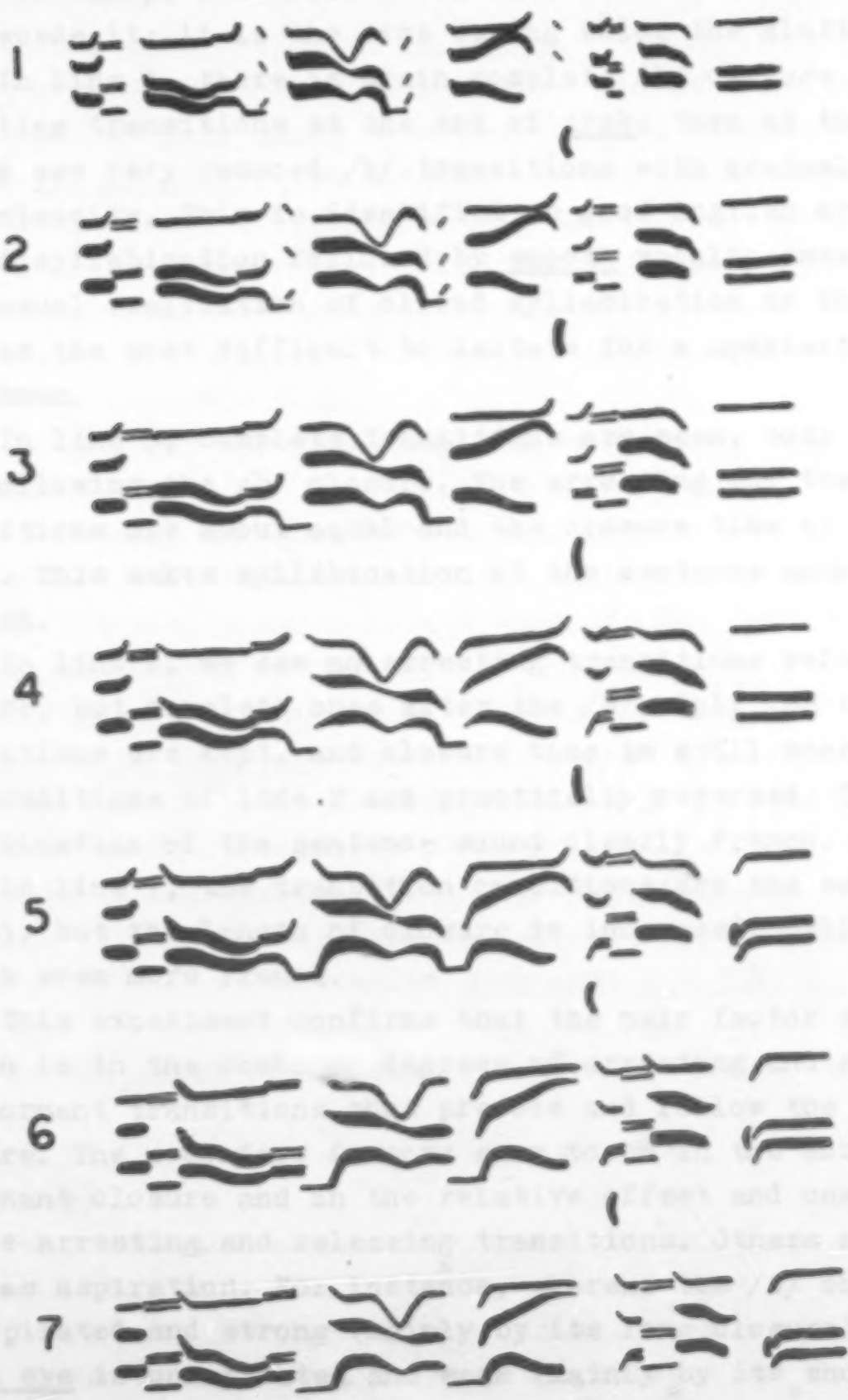


Figure 7

In the third line, the closure gaps are decreased but the onsets are still brusque. This does not sound English, or German or anything else, meaning that in English or German, if the onset is sharp, a sufficient closure (silence) time is required to precede it; it is the time during which the glottis is closed.

In line 4, there is again complete /b/ closure with maximal arresting transitions at the end of Arab; then at the onset of ate we see very reduced /b/ transitions with gradually increasing intensity. This is identified as good English or German closed syllabication followed by smooth vocalic onset -- the most usual realization of closed syllabication in English, as well as the most difficult to imitate for a Spaniard or a Frenchman.

In line 5, complete transitions are seen, both preceding and following the /b/ closure. The arresting and the releasing transitions are about equal and the closure time is relatively short. This makes syllabication of the sentence sound rather Spanish.

In line 6, we see no arresting transitions before the /b/ closure, but complete ones after the /b/. Only the releasing transitions are kept, and closure time is still short. Here the conditions of line 2 are practically reversed. This makes syllabication of the sentence sound clearly French.

In line 7, the transition conditions are the same as in line 5, but the length of closure is increased. Syllabication sounds even more French.

This experiment confirms that the main factor of syllabication is in the combined degrees of arresting and releasing of the formant transitions that precede and follow the consonant closure. The secondary factors seem to be in the duration of consonant closure and in the relative offset and onset intensities of the arresting and releasing transitions. Others may appear, such as aspiration. For instance, whereas the /t/ of gray tie is aspirated and strong (mainly by its long closure), the /t/ of great eye is unaspirated and weak (mainly by its short closure).

only a reduction of intensity.

Although this experiment is but a first step in a long series, a tentative comparison of syllabication for the type ab a (an aim, zum einen, d'une autre, las aves) among the four languages might be made here.

In English, maximally arresting consonant transitions always precede the closure. The onset of the following vowel can be either very sharp with glottal stop closure and transitions, fairly sharp with glottal closure but no marked glottal stop transitions, or smooth with no glottal closure, much reduced ($1/3$ of arresting) releasing consonant transitions, and very gradual intensity increase.

German has the same three types but uses the very sharp onset more frequently.

In Spanish, the arresting and releasing transitions are somewhat equal.

In French, the arresting transitions are very weak (from zero to $1/3$ of releasing), and the releasing transitions are maximal. Closure time is longer than for the "smooth" type in English or for the normal type in Spanish.

Further experiments will include comparison of intervocalic consonants in the word, where differences are of the same type as between words but much more subtle and, in addition, more dependent on the place of stress.

Next in importance, for comparative teaching purposes, is the case of phonetic realization of internal juncture for a sequence of two vowels, as in stay out, die Uhr, là aussi, una isla.

Here Spanish is markedly different from the three other languages.

In English, the second vowel may have a sharp onset or a smooth one.

In German the sharp onset, with glottal stop, is much more frequent than in English, but the smooth onset is also heard.

In French, the smooth onset of the second vowel is the rule. There is no voice interruption between the two vowels, only a reduction of intensity.

In Spanish the smooth onset is also the rule, but, in addition, the first vowel, if unstressed and not closer than the second, tends to fall (sinalefa): busc(a) una, cuart(o) oscuro; if unstressed and closer than the second, it tends to become a semi-vowel: ni una vez, su amigo.

The acoustic and articulatory features of this type of juncture will be studied experimentally and statistically.

X. SYLLABLE TYPES

As a preliminary comparison of the characteristics of syllables, we have gathered data on 1000 contiguous syllables in a narrative text for each of the four languages.

The ten most frequent syllables that emerge in each language are obviously too dependent on common words, yet they must be somewhat representative of the syllabic impression one gets as he hears a language.

1. The ten most frequent syllables are:	Average No. of sounds <u>per syllable</u>
in English: e, ev, ðe, in, ænd, aɪ, tu, ri, for, ʃæt	2.2
in German : unt, er, tɐ, gə, das, tsu, di, der, zai, ain	2.6
in French : te, de, le, se, də, la, e, ma, re, a	1.8
in Spanish: a, de, e, la, ʎa, i, ke, to, na, ra	1.7

2. Their load, measured by the number of consonants per syllable, is in the following proportions:

German:	1.6
English:	1.2
French:	.8
Spanish:	.7

This syllabic load is likely to correlate with syllabic speed, or rate of syllables per second, a feature we shall also compare. The more loaded the syllables, the slower the rate should be and the heavier the subjective impression.

3. The proportion of closed, or checked, syllables (mar-tes) to open, or free, syllables (ca-mi-no), is:

	<u>Closed</u>	<u>Open</u>
German	64.1%	35.9%
English	58.0	42.0
Spanish	28.3	71.7
French	23.5	78.5

4. The proportion of Consonant-Vowel-Consonant, Vowel-Consonant, Consonant-Vowel, and Consonant-Consonant-Vowel syllables is:

	CVC	VC	CV	CCV
German	38.8%	10.3%	26.3%	5.5%
English	28.0	12.8	29.8	4.9
Spanish	20.1	3.0	54.7	9.3
French	16.4	.8	55.3	16.4

XI. TENSION

Some languages are said to be more tense, or more lax, than others. What is probably meant is that the muscles of articulation are more contracted, and/or that more muscular energy is spent per unit of time or per sound. This is not easy to evaluate. Some measurable features of tension must be found. Upon the accepted notion that, historically, in the evolution of languages we know, laxness of articulation produced diphthongization and affrication, which correspond to unsteadiness of holds and slowness of movements, we will attempt to measure tension by the steadiness of closed and open positions and the speed of

closing and opening movements that relate to the perception of consonants. However, we will assume that the presence of diphthongs and affricates in a language is not necessarily a mark of present laxness but a remnant of past laxness. Present diphthongization of monophthongs and affrication of stops is the real mark of laxness. Other means of measuring tension will also be sought.

COMPARING THE VOCALIC FEATURES OF ENGLISH, GERMAN, SPANISH AND FRENCH¹

This is the third in a series of articles on the general phonetic characteristics of American English as compared with those of German, Spanish and French. Following a description of the research technique we have developed (IRAL, 1-2:85-97, 1963), and which is four-fold -- radiographic observation, spectrographic analysis, spectrographic synthesis and statistical survey -- the prosodic features of those four languages were compared (IRAL, 1-4, 1963). Now the phonetic features that characterize the vowels of each of the four languages will be presented, with emphasis on what distinguishes one language from the others.

Our comparative study of vowels is in progress under the following subdivisions: 1. Articulatory description. 2. Acoustic description. 3. New vowel sounds. 4. Vowel distribution. 5. Vowel frequency of occurrence. 6. Vowel duration system. 7. Neutral position. 8. Loss of vowel color. 9. Effect of consonant anticipation. 10. Diphthongization. 11. Vowel color and syllable type. 12. Vowel attack and release. 13. Vowel nasality. Present results are only partial, naturally, but it is hoped that even where the research is the least advanced, an understanding of the direction it is taking will benefit the language teacher.

1. ARTICULATORY DESCRIPTION OF VOWELS

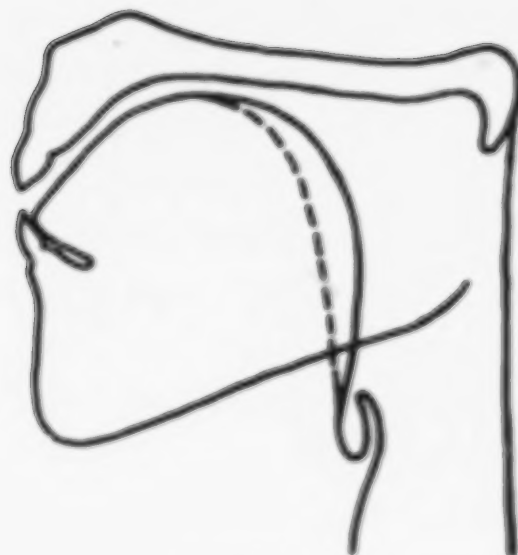
The study of tongue position and of front and back cavity shape or dimensions by means of X-ray still photography has proved to be largely misleading. The necessity of posing may introduce gross errors -- lack of naturalness, lack of realism are indicated by inconsistent tongue positions for sounds pronounced by the same subject at different intervals of time.

¹ The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

In order to discover objectively the tongue position (or the tongue motion) that is most typical of a language for a given sound, it is necessary to observe the tongue in motion in the course of natural speech by means of motion picture X-ray. Moreover, it is important to place the vowels that are being compared in similar phonetic environment and prosodic conditions. With these requisites in mind we have taken motion-picture X-rays of some eight pages of comparative material of the following type. To compare English /i/ to German /i/, Spanish /i/, or French /i/, pairs of sentences like these are used: Ein Dieb war er nie/It's deep to his knee; La lista esta aqui/At least it's a key; Chez qui sont partis ses habits/McKee has a bee in his tea. It is only from an appropriate variety of such material in motion picture that the characteristic vowel positions and movements can emerge. Figure 1 presents the typical /i/ vowel profiles, obtained by this method, in the four languages. Since the English /i/ is somewhat diphthongized, the tongue being in constant motion while the sound changes from [ɪ] to [i] to [j] (see Part 10 on Diphthongization), we use here a tongue position near the end of [i] before [j] is reached. Thus, compared with French, German and Spanish /i/ vowels, English /i/ is higher at mid-tongue, less fronted at the blade (the front cavity has a typical cone shape) and slightly more retracted toward the pharyngeal wall; the jaws are more open, the lips more open and less retracted at the corners. A large front cavity and a smaller back cavity should result in a lower second formant and a higher first formant, respectively. This will be confirmed in the next chapter on the acoustic description of vowels. Differences among the three target languages can also be noted: Spanish is less close and less fronted than French -- a step in the direction of [e] -- but with similarity of shape. German is only less fronted, with a tendency toward the cone shape of English for the front cavity.



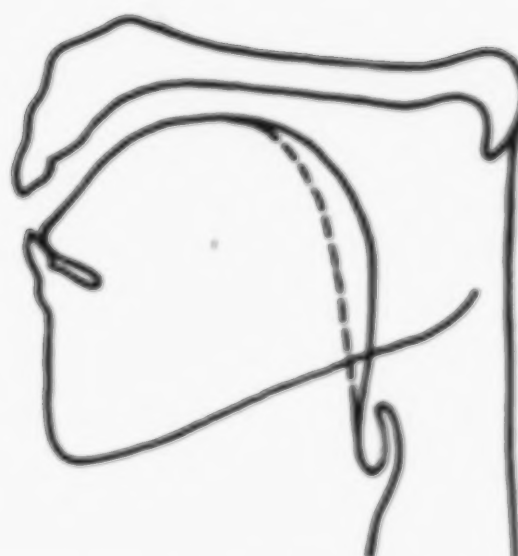
ENGLISH



FRENCH



GERMAN



SPANISH

Figure 1

2. ACOUSTIC DESCRIPTION OF VOWELS

An even more objective way of comparing the vowels of languages than X-ray observation is furnished to us by recent advances in acoustic research. It consists in making "acoustic charts" of the vowels of a language directly from the sound wave of speech as it is made visible by spectrographic analysis and synthesis. Such acoustic charts can easily be understood as their shapes are remindful of the "articulatory charts" on which dots represent the highest point on the tongue profile. On the articulatory chart of Figure 2, the tongue is high-front for [i], high-back for [u], low-front for [æ], and low-back for [ɑ].

The acoustic chart condenses more information than the articulatory chart, because it is not limited to tongue position; it takes into account all other articulators, including, for instance, the lips, which by rounding can change [i] to [y] without moving the tongue.

Vowels are musical chords, mainly characterized by the frequency (note) of their two lowest formants, Formant 1 and Formant 2. (The only exceptions are /ɜ/ and the nasal vowels which require three formants.) The frequencies of Formants 1 and 2 are listed in Table 1 and are presented on a musical scale in Figure 3. Those formants reflect the notes of resonance of the vocal tract cavities. Roughly, we can think of Formant 1 as having the frequency of the back cavity (pharynx) and Formant 2 as having the frequency of the front cavity (mouth). Thus, when we measure the frequency of Formants 1 and 2 on spectrograms, it is somewhat as if we measured the resonance notes of the back and front cavities. Obviously the shape and size of those two cavities are different for each vowel, and as a result the musical chords are also different.

All charts in Figures 4, 5, 6 are acoustic charts obtained by plotting the frequencies of Formants 1 and 2 in abscissa and ordinate. For instance, [i] is to the left and high on the charts because it has a very low frequency in Formant 1 and a very high frequency in Formant 2. In articulatory terms this

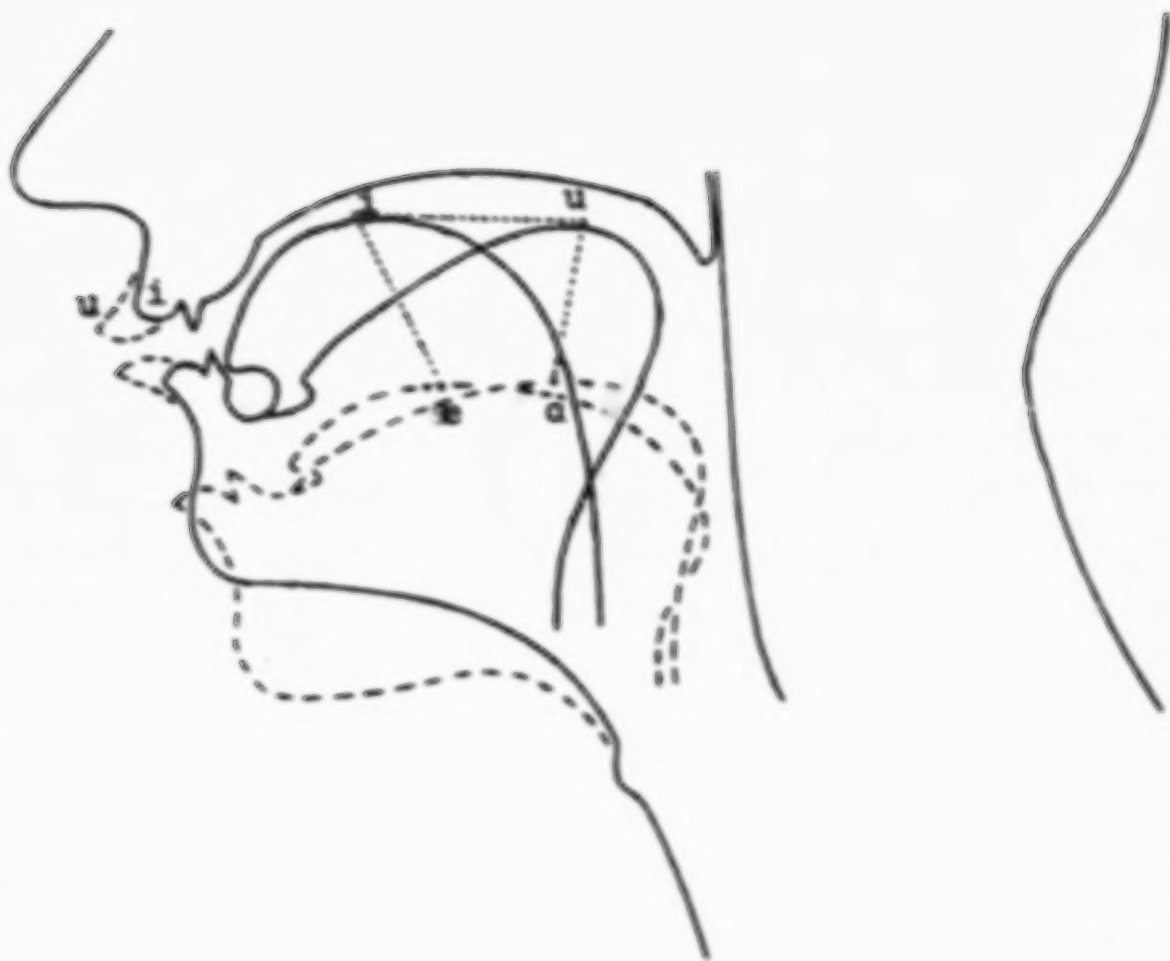


Figure 2

Table 1

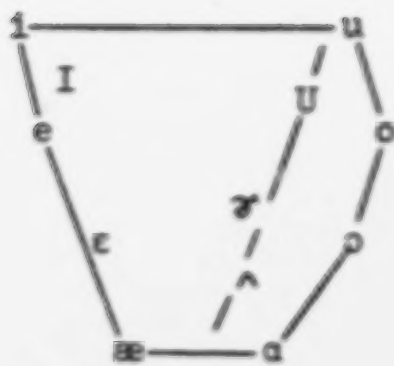
VOWEL FORMANT FREQUENCIES

<u>Vowel</u>	<u>English</u>	<u>French</u>	<u>German</u>	<u>Spanish</u>	
/ɪ/	2250 300	2500 250	2250 275	2300 275	F2 F1
/ʏ/		1800 250	1750 275		
/ɪ/	1950 350		2100 325		
/ʏ/			1700 325		
/e/	2100 400	2200 375	2050 375	1900 450	
/ə/		1600 375	1600 375		
/ɛ/	1800 550	1800 550	1900 500		
/a/, /ʌ/	1200 600	1400 550	1550 500		
/æ/, /a/	1700 750	1350 750	1250 750	1300 725	
/ɑ/	1100 750	1200 750	1150 750		
/ɔ/	900 550	950 550	900 500		
/o/	800 400	750 375	850 375	900 450	
/ʊ/	1000 375		950 325		
/u/	900 300	750 250	850 275	800 275	
/ɐ/		750 600 (250)			
/ɑ/		950 600 (250)			
/a/		1350 600 (250)			
/ɛ/		1750 600 (250)			
/ɜ/	(1500) F3 1200 500				

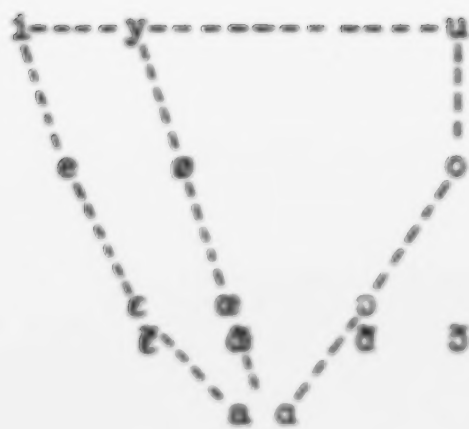
Figure 3

VOWEL FORMANTS

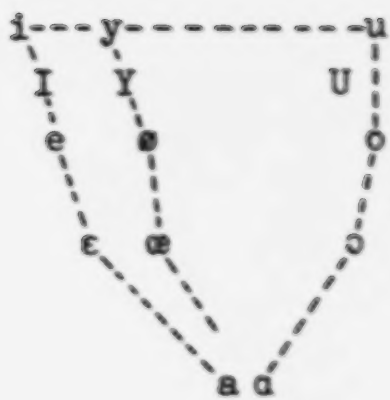




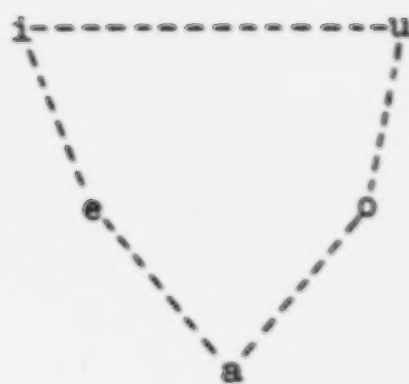
English



French

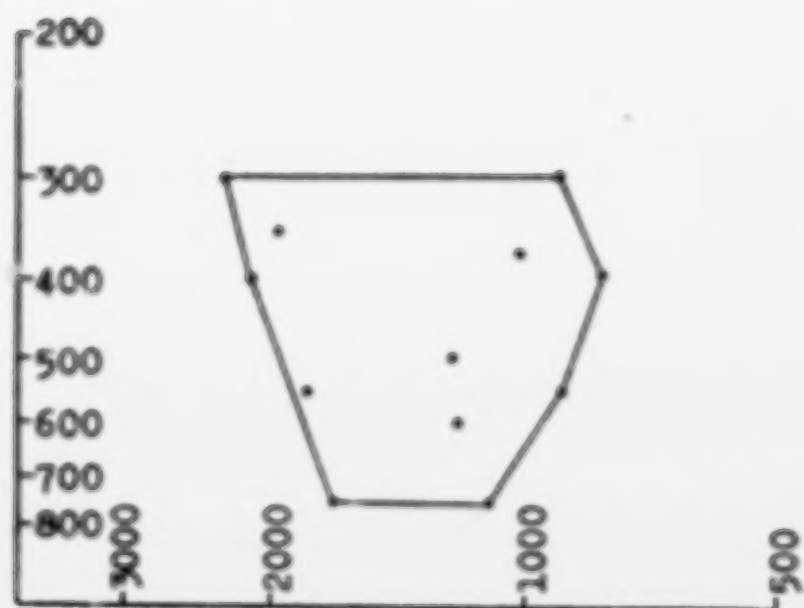


German

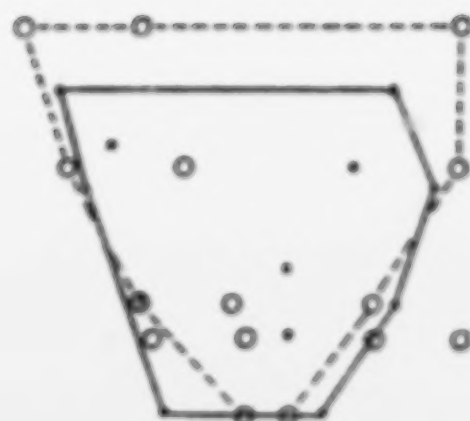


Spanish

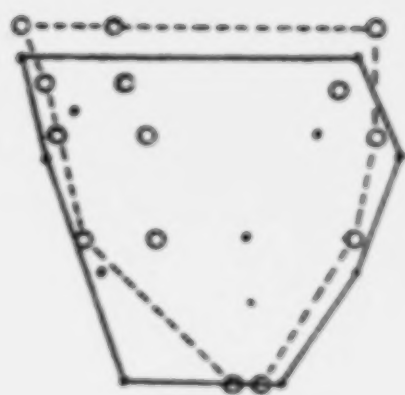
Figure 4



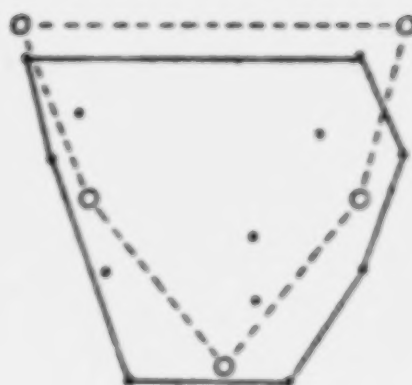
English alone



English with French

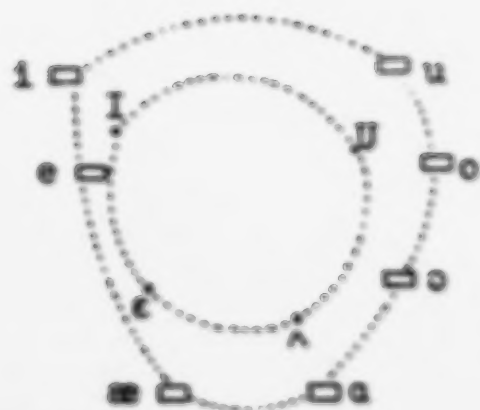


English with German

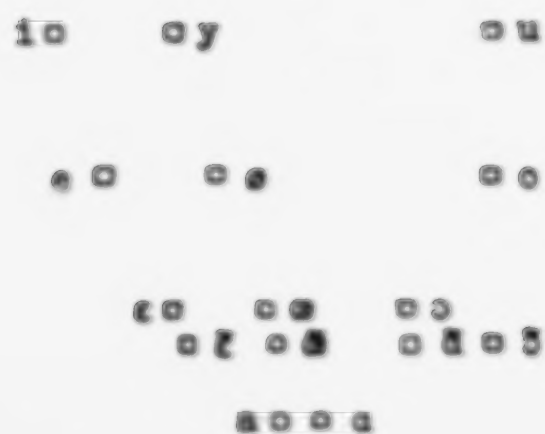


English with Spanish

Figure 5



English



French

②



German



Spanish

Figure 6

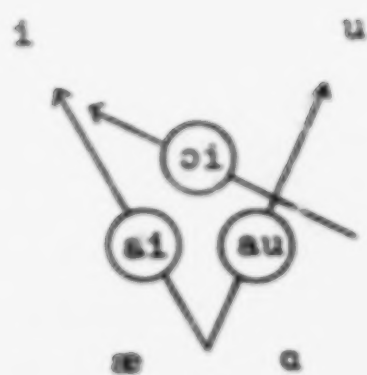
means that [i] has a very large cavity in back and a very small cavity with spread-lip opening in front (see Figure 2). If the opening of the front cavity were closer because of lip rounding, as for [y], the note of Formant 2 would be lower -- less toward the left on the chart.

Methods of frequency measurement for all vowels. Linguists from many countries have attempted to measure Formants 1 and 2 on spectrograms and have published their results. When we verify those results by synthesis, they do not stand the test, either because of technical faults in measuring or because of variations due to human elements not entirely explained -- the formants of a woman can be about 15 percent higher than those of a man with the same dialect, and even persons of the same sex can show considerable divergence of formant frequencies for a similar sound.

The frequency ratio among formants by different persons is fairly constant, however, so that spectrographic analysis can be used with confidence to obtain the relative values of formant frequencies. For instance, the second formant is always lower for [y] than for [i], regardless of the speaker's sex or age. But in order to compare the vowels of different languages, absolute values according to a uniform standard were needed.

This problem was solved by using synthetic (artificial) vowels made on our speech synthesizer at a uniform fundamental frequency of 120 cycles per second, which is a good standard of fundamental frequency since it corresponds to a typical average male voice. Formant frequencies of synthetic vowels produced from hand-painted spectrograms were varied, and the resulting vowels were tested by natives of each language, and by linguists, until found typical. Thus the values obtained for Formants 1 and 2 of all vowels on Figures 3 and 7 are comparable among themselves and comparable from one language to another -- with the understanding that they are valid for a male voice speaking at a fundamental frequency of 120 cps.

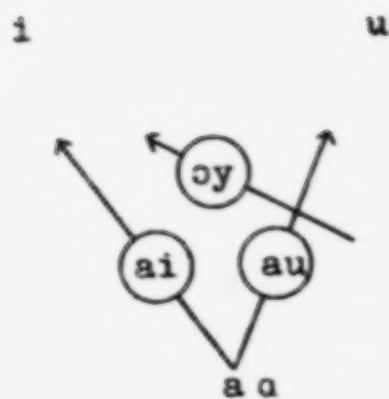
Moreover, the frequencies of all vowels on Figures 3 to 7



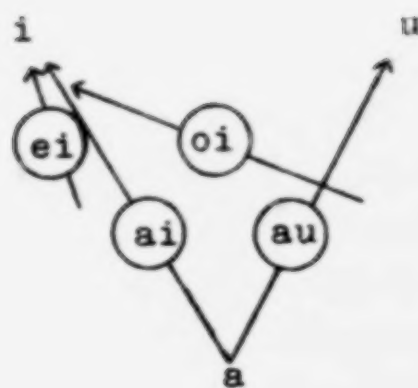
English

None

French



German



Spanish

Figure 7

are the most characteristic single frequencies, those by which a vowel is best identified by natives, in the absence, or regardless, of diphthongization.

Even though our vowel charts have been obtained by acoustic measurements, we may use the articulatory terminology everyone is used to in discussing them.

Let us now use Figures 4, 5, 6, and 7 to compare the English vowels to those of the three target languages, first as a whole, then individually. ("English" will always mean "American English.")

1. The vocalic system of English is rich and complex: a minimum of 12 monophthongs contrasting among themselves by color (formant frequencies) and by length (Figures 4, 5, 6), and 3 diphthongs (Figure 7). It is much richer than Spanish: 5 monophthongs and 4 main diphthongs, and only slightly less so than German: 14 monophthongs and 3 diphthongs; or French: 15 monophthongs and no diphthongs. (The problem of whether the two mid-series [e, ø, o], [ɛ, œ, ɔ] represent 3 or 6 phonemes in French is unsolved. They are in complementary distribution in some series: nez-nette, peut-peuvent, mot-mort, which have a high return, but they are in contrast in others: pré-prêt, jeûne-jeune, paume-pomme, which have a low return.)

2. The vocalic system of English is more open (low) than those of the three target languages. Its close vowels are less close. Its center of gravity is lower. And its low vowels are more extreme (close to cardinal vowels) than its high vowels, which is not the case with the target languages.

3. If we call "back" vowels all those that are clearly to the right of center on the acoustic charts, the vocalic system of English appears as richer in back vowels than the three target languages. Figure 4 shows that English has two back and one front series; in French and German the proportion is inverse: two front and one back series; and in Spanish the front and back vowels are in equal ratio.

4. The vocalic system of English is less rounded than those of the three target languages. Not only is the ratio of rounded

vowels smaller ($1/3$ in English, $1/2$ in Spanish, $2/3$ in French, $2/3$ in German), but those vowels which use lip rounding in English, such as /u, o, ɔ/, have the rounding less pronounced than corresponding vowels in French, German and Spanish.

Indeed, one of the most characteristic features of English is its somewhat back and unrounded series of vowels /ʌ, ɜ, ʊ/ and perhaps even /u/. None of the three target languages possesses such a series.

5. High-close and high-open vowels /i, ɪ, u, ʊ/ are less extreme (not so near the outside on the acoustic charts) in English than in the three target languages. The greatest difference is with French in /i, u/, but it is also noticeable with German in /i, ɪ, u, ʊ/, and with Spanish in /i, u/. In other words, the high-close and high-open vowels of English tend slightly toward center, unlike the corresponding vowels in French, German and Spanish.

On the other hand, low vowels like /æ/, /ɑ/ are more extreme (more peripheral on the acoustic charts) in English than in the target languages.

"Centering" further appears as a distinguishing characteristic of English if we note that about 90 percent of unstressed vowels turn to some sort of schwa (neutral vowel) and are located inside the inner circle of the English chart of Figure 6. Since unstressed vowels comprise about 61 percent of all English vowels in connected speech, the proportion of vowels located in this inner circle is over 50 percent. The proportion of centered vowels is very much smaller in the three target languages. (The effect of unstressing on vowel color is the subject of another chapter.)

6. In the English system, vowel duration plays a distinctive role. Four vowels, /ɪ, ʊ, ɛ, ʌ/, are shorter than the others, on the average; and those shorter vowels are all nearer the center of the vowel chart. As shown by the two dotted-line circles in Figure 6, the longer vowels are located on a larger peripheral circle; the shorter ones on a more central circle. Briefly stated, in English, vowel shortening correlates with vowel center-

ing, for stressed as well as for unstressed vowels. (Historically, shortening preceded centering and may be assumed to have caused it.) In German, where differences of length are more pronounced than in English, the correlation is somewhat different: a) long and short vowels go in pairs; b) except for short /a/, the shorter vowels are more open and slightly more centered than the corresponding long ones. In Spanish and French, duration has no stable contrastive function. (The subject of vowel duration is treated in another chapter.)

7. In most American English dialects, all vowels are more or less diphthongized. This diphthongization plays an important role in distinguishing vowels that seem too close on the chart. For instance /ɪ/ diphthongizes toward /ə/. This helps in distinguishing it from /e/ which diphthongizes toward /i/. The three target languages do not noticeably diphthongize, not even German in its long /e/ and /o/. (Diphthongization is the subject of a special chapter.)

8. Let us summarize the phonetic characteristics of each vowel system. Comparatively, English vowels are predominantly low, back, unrounded, with a strong tendency to center the short and unstressed (except when very low). Duration contributes to vowel distinctions. All English vowels are more or less diphthongized. Most characteristic series: the somewhat back-unrounded /u, ʊ, ɜ, ʌ/.

9. Comparatively, French vowels are predominantly high, fronted, rounded, and extreme (except when low). The /i, e, y, ɛ, u, o/ are nearly cardinal. The open/close difference in the mid-vowels is very pronounced. Duration is negligible in vowel distinction. There is no diphthongization. Most characteristic series: the front-rounded /y, ɛ, œ/ and the nasals /ɛ̃, ɔ̃, ɪ̃, ʊ̃/.

10. Comparatively, German vowels are predominantly high, fronted, rounded, more extreme than English but less than French. Differences of duration play an important role in vowel distinction (more than in English). For every long vowel except /a/ there is a corresponding short one which is more open. There is

no diphthongization, not even in long, close /e/ and long, close /o/. Most characteristic series: the front rounded /y, Y, ø, œ/, which is different from the French series in that it offers four degrees of opening within a small range; French offers three degrees within a larger range. Differences of duration help in distinguishing vowels that may be somewhat close on the vowel chart, such as /y/ from /Y/, or /Y/ from /ø/.

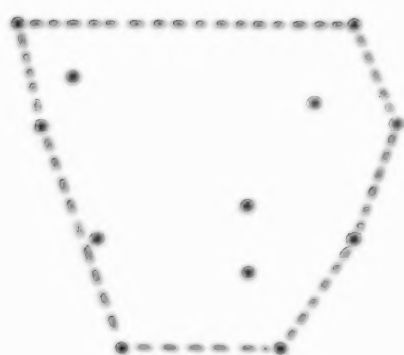
We are aware that long, open /ɛ/ exists as a phoneme for some speakers of North German. It is not included on our vowel charts on the ground that a majority of speakers, and of our informants, tend to assimilate it to the long, close /e/ phoneme.

11. Comparatively, Spanish vowels are predominantly high and peripheral. They are distant from each other. The open/close difference of mid-vowels is negligible. Differences of duration are not distinctive. There is no diphthongization. Most characteristic series: the mid-open/mid-close vowels /e/ and /o/, which are midway between the French and English close and open vowels.

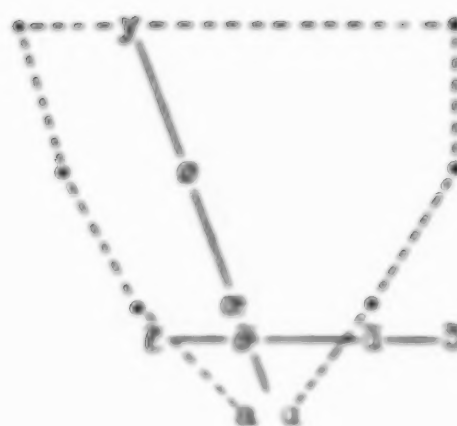
3. NEW VOWEL SOUNDS

We shall consider here the vowel phonemes that are radically different from those of another vowel system (see Figure 6).

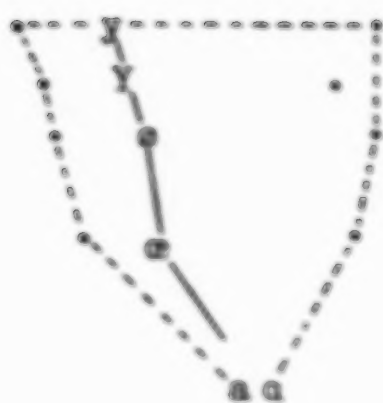
1. An American learning French has to meet three new sets of vowels in the following order of difficulty: a) The three front-rounded vowels /y/, /ø/, /œ/. Since he is used to combining rounding with tongue-backing, he has to acquire the very new habit of rounding while fronting his tongue: u/y, o/ø, œ/œ: sous/su, soit/ceux, sol/seul. b) The two low a sounds, which are very close to each other and located between the two American low sounds on the acoustic chart. The /ɑ/ has a much lower frequency of occurrence than the /a/ (about 150 lexical words vs 6000), but minimal pairs such as tache/tâche, pâtte/pâte are still alive in Paris. c) The four nasal vowels: /ɛ̃/, /ɔ̃/, /œ̃/, /ɑ̃/. Note on Figure 6 that acoustic analysis has clearly demonstrated that those nasal vowels do not have the same place of articulation as their (supposedly) oral counterparts.



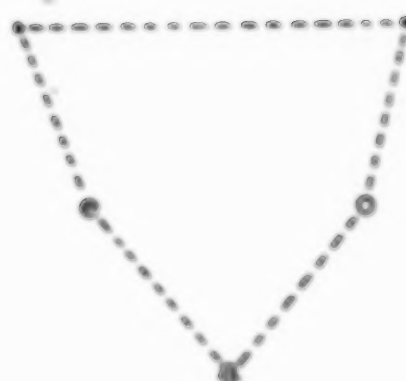
English



French



German



Spanish

Figure 6

2. An American learning German has to face two new groups of vowels: a) The four front-rounded vowels, long /y/, short /Y/, long /ø/, short /ʊ/. As in French, he has to learn to combine lip-rounding with tongue-fronting, not with tongue-backing: u/y, ʊ/Y, o/ø, ɔ/ɔ: Kuhle/kÜhle, Mutte/Hütte, Boden/Böden, Holle/Hölle. These vowels should be more difficult to distinguish than the French ones since there are four of them in less articulatory space than for three in French. However, as mentioned above, their distinction is considerably aided by the long/short distinction, which is much more marked in German than in English. b) The two a sounds, which are farther back than the French ones, but not so far back as the American /ɑ/. (Phonemic symbols have a different value in each language, of course.)

3. An American learning Spanish has to face three new vowels that are midway between pairs of corresponding American vowels: a) A pair of vowels (symbols: /e/, /o/) that are less close than American /e/, /o/ and less open than American /ɛ/, /ɔ/. b) An /a/ that is clearly less fronted than American /æ/ and less backed than American /ɑ/.

Summary of new vowel sounds for an American

In French: y ø œ, a ɑ, ɛ ɔ ɛ ɔ

In German: y Y ø œ, a ɑ

In Spanish: e o a

4. VOWEL DISTRIBUTION

It is important for a language teacher to know in which positions vowels occur in the target language as well as in the native language of the students. As a start toward complete pictures of positional and allophonic distribution, Tables 2, 3, 4 and 5 present, in examples, the positional distribution of stressed vowels and the schwa in English, German, French and Spanish.

In American English, (Table 2), the short vowels do not occur in final stressed position. Neither does the /ə/, which is not a short vowel in American English but is in British English. In addition, the short /ʊ/ does not occur initially.

Table 2

English Stressed Vowels

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
i	eel	seen	see
I	if	sit	
e	ate	bait	day
ɛ	ebb	sell	
æ	at	cat	
a	are	clock	blah
o	ought	thought	saw
o	own	note	no
U		book	
u	ooze	boot	two
ʌ	up	sun	
ɜ	earth	first	fur
aɪ	ice	time	pie
au	out	down	cow
ɔɪ	oil	boil	boy
(ə	above	believe	sofa)

Table 3

German Stressed Vowels

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
a	Aale	Kahn	da
a	alle	kann	
e	Ehre	Segen	See
ɛ	Ernte	Säcke	
i	ihn	bieten	wie
I	in	bitten	
o	Ofen	Sohn	so
ɔ	offen	Sonne	
u	Uhrzeit	Ruhm	Kuh
U	Unzeit	Rum	
ø	Öfen	Höhle	Bö
œ	öffnen	Hölle	
y	üben	Mühle	früh
Y	üppig	Müller	
ai	Eimer	nein	zwei
au	auch	Raum	schlau
ɔy	Europa	neun	neu
(ə		arbeitet	eine)

Table 4

French Stressed Vowels

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
i	file	mille	lit
e	et		pré
ɛ	elle	belle	(prêt)+
a	acte	patte	là
ɑ	âme	pâte	las
o	haut	(paume)	dos
ɔ	homme	pomme	
u	outré	foule	tout
y	une	mule	vu
ø	eux	(jeune)	deux
œ	heure	jeune	
ɛ̃	Inde	quinze	pain
ɑ̃	humble	lundi	parfum
ɑ̃	anse	sens	cent
ɔ̃	onze	bombe	son
(ə)		premier	)

+ Those in parenthesis are unstable or of very low frequency.

Table 5

Spanish Stressed Vowels

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
i	ir	fin	dí
e	el	ver	fe
a	has	tan	da
o	hoz	dos	no
u	un	luz	tú
au	auto	causa	jau
ei		reina	ley
ai	aire	baile	fray
oi	oiga	lloica	doy

In German (Table 3), as in English, the short vowels do not occur in final stressed position.

In French (Table 4), the tendency is clearly for mid-close vowels not to occur medially and for mid-open vowels not to occur finally, in accordance with the Law of Position ("In an open syllable mid-vowels are close: mes, meut, mot [me, mɔ, mo]; in a closed syllable mid-vowels are open: mère, meurt, mort [mɛr, mɛr, mɔr]"). But the influences of spelling or etymology have created a few exceptions, so that in fact, of the mid-close vowels, only [e] never occurs medially -- [o] and [ø] occur medially in a few hundred words, mainly of the types paume [po:m], pose [po:z], jeûne [ʒø:n], creuse [krø:z], with concomitant extra length; and of the mid-open vowels, only [ɔ] and [œ] never occur finally -- [ɛ] occurs finally in the dialect of nearly one-third of northern French speakers, mainly in words of the types prêt [prɛ] and j'allais [ʒalɛ]. (Southern France observes the Law of Position without exception. But in the unaffected speech of the most cultivated class, [pre], [ʒalɛ] is not only the pronunciation of all southern speakers but also that of a majority of northern French speakers.)

Spanish shows all vowels in all three positions except for the diphthong /ei/ which does not occur initially.

Summary: The missing positions are related, in German and English to the short/long differences, in French to the Law of Position.

The schwa vowel [ə] behaves differently in each of the four languages. It occurs in three positions in English, two in German, one in French, and does not occur in Spanish, where natives use an [e] sound to mark their hesitation.

5. VOWEL FREQUENCY

Frequency of occurrence of vowel phonemes is a good indication of the auditory impression a foreigner receives of a language when he enters the country in which it is spoken. This impression should be recreated for the new student of a target language

with the help of appropriate material in which the sounds would be balanced according to frequency of occurrence. The first learning sequences should also take into account such frequency of occurrence.

Table 6 presents a preliminary comparison of vowel phoneme frequency based on short samplings of combined narrative and spoken material. (This will be done on a large scale later.)

Stressed and unstressed vowels are not separated in this table. In German, any vowel can occur in stressed as well as in unstressed position except /ə/, which is always unstressed. In the three other languages all vowels can occur stressed as well as unstressed. (Loss of vowel color under lack of stress is mentioned in another chapter.)

German frequency of occurrence has much in common with English. Both have /ə/ as their most frequent vowel. (However, in German /ə/ is always unstressed, whereas in English it occurs as stressed in a proportion of approximately 1 to 8: one [ʌ] to eight [ə] sounds.) Furthermore, the five most frequent vowels are /ə, ɪ, æ, ɛ, e/ in English, and /ə, a, e, ɪ, ɛ/ in German. The three diphthongs have rather low frequency in both languages and their frequency order is the same: /ai/, /au/, /ɔi/. We find this to be even more true in Spanish, where the four diphthongs /ei/, /ai/, /au/, /ɔi/ occupy the last four positions on the frequency list.

German and French have in common the front-rounded series, the [y], [ø] family, which have no equivalent in English nor in Spanish. However, in both French and German, those front-rounded vowels have a very low frequency of occurrence. They serve therefore much less than is generally assumed in characterizing the auditory impression of those languages, and should not be taught too early.

The same thing must be said of the French nasal vowels. They have a much lower frequency of occurrence than the corresponding oral vowels: /ã/ is lower than /a/ and /ɑ/, /ɛ̃/ than /e/ and /ɛ/, /ɔ̃/ than /o/ and /ɔ/, /œ̃/ than /ø/ and /œ/. They also have a lower frequency than the oral vowels from which they mainly

Table 6

Comparative Frequency of Occurrence of Phonemes

English		French		German		Spanish	
Vowel	%	Vowel	%	Vowel	%	Vowel	%
ə	22.99	e	19.29	ə	23.88	e	32.80
ɪ	14.44	a	16.69	ɪ	11.52	a	30.38
æ	9.44	i	12.39	a	10.72	o	21.56
ɪ	8.49	ə	7.60	ai	7.88	ɪ	10.48
ɑ	6.99	ɑ̃	7.59	ɛ	7.08	u	4.23
ɛ	6.85	ɛ	6.69	e	6.73	eɪ	0.20
u	5.60	u	6.40	ɪ	5.88	ai	0.15
ai	5.50	ɔ	5.05	u	5.28	au	0.15
o	4.95	y	4.70	ɑ	3.89	oi	0.05
ɜ	4.40	ʒ	3.85	u	3.84		
e	3.95	o	2.60	ɔ	3.74		
au	2.20	ɛ̃	2.45	au	3.14		
ɔ	2.00	œ	1.80	o	2.29		
ʊ	2.00	ø	1.70	y	1.30		
oi	0.20	œ̃	1.05	Y	0.98		
		ɑ	0.15	ø	0.95		
				ɔy	0.55		
				œ	0.35		
	<u>100.00</u>		<u>100.00</u>		<u>100.00</u>		<u>100.00</u>

came: /a/, /e/, /ɛ/ for /ā/; /i/ for /ē/; /y/ for /ÿ/; and /u/, /o/, /ɔ/ for /ū/. (Denasalization toward the end of the Middle Ages can account for this low frequency of nasals.) On the basis of the low frequency of occurrence of French nasal vowels, French is a more oral language than appears on the surface, and the nasals should be taught late in the period of conditioning to new sounds.

Spanish and French have in common a high proportion of fronted vowels. They share their first two vowels: /e/ and /a/ have the highest frequency of occurrence in both languages.

Finally, Spanish and French also have in common a high proportion of vowels (against consonants): 43.6% and 43.2%, as opposed to 37.4% and 36.3% in English and German, respectively.

6. VOWEL DURATION SYSTEMS

Variations in vowel duration can be distinctive (contribute to change in meaning) or non-distinctive. In English, the difference of duration between bit and beat, in a ratio of nearly 2 to 3, about 10 centiseconds and 14 centiseconds in equally stressed positions, is distinctive. It has to be learned when learning English, and one easily becomes conscious of it. But the differences between bit (10 cs) and bill (14 cs), bit and bid (16 cs), bit and bin (18 cs), are non-distinctive and made unconsciously. They are conditioned by the force of articulation of the following consonant. German and French also have both learned (distinctive) and conditioned (non-distinctive) variations in vowel length, but Spanish only has conditioned ones. Let us consider the two types separately.

A. Distinctive vowel length.

a) The importance of distinctive vowel length is greatest in German, then in English, then in French.

b) In stressed position the ratio of short vowel to long vowel counterpart is nearly

2 to 5 in German	bitte/biete
2 to 3 in English	bit/beat
3 to 4 in French	mettre/maitre

c) In German the role of duration in the i/i type distinction is considerable in stressed position. It is perhaps greater than the difference in color. In English the role of duration in the i/i type distinction is certain but not considerable. It is probably smaller than the difference in color. In French the role of duration in the e/ε type distinction (mettre/maître) is strong when the difference of duration is observed.

d) In German and English, this difference of length in stressed position is in a constant ratio and is always observed. In French it is in an unstable ratio and is made irregularly. It is observed only in careful diction and by a minority of speakers.

e) In German, the distinctive difference of vowel duration in stressed position applies to seven pairs of vowels: i/i, y/Y, u/u, e/ε, ø/œ, o/o, a/a (Refer to Figure 6). For every long vowel there is a corresponding short one. Duration differences are integrated in the system of vowel phonemes. In American English, distinctive duration is much less integrated in the system. There are three pairs in which duration plays some distinctive role: i/i, u/u, e/ε; but there is no one-to-one correspondence among the long and the short as a whole. It is perhaps more realistic to say that seven vowels are longer: /i, u, e, o, œ, a, ɔ/, and that four are shorter: /ɪ, ʊ, ε, ʌ/. (The diphthongs are long, naturally, both in German and in English.)

In French the learned distinction short/long applies to two pairs of vowels, but, as we said, it is very unstable. The pairs are:

e/ε: mettre/maître, tette/tête, etc.

a/a: tache/tâche, patte/pâte, etc.

Because of the instability of these duration differences, it is often said that duration is not distinctive in French. In two other types of pairs, a difference of duration is made regularly, but it seems to be conditioned rather than learned. These types include six pairs of vowels:

ɔ/o:	pomme/paume	open and close vowels in closed syllable
œ/e:	jeune/jedne	
ɛ/ɛ:	messe/mince	oral and nasal vowels in close syllable
ɔ/ɔ:	mode/monde	
a/ɑ:	bac/banque	
œ/œ:	au neuble/homme humble	

f) Note. The assumption that, in English, shortening tends to center a vowel is confirmed by the great number of vowels that turned to schwa, under the influence of Celtic substratum, when unstressed and short. Compare politics and political: [palet:ks], [pol:tekel].

B. Non-distinctive vowel length.

The length of vowels can also vary under the influence of a great many physiological factors. Some are found in the vowels themselves: diphthongs are on the average longer than monophthongs; open vowels are longer than closer ones (this has been observed in English, German and Spanish); nasal vowels are longer than oral ones (this has been observed in French). Other factors are found in the contiguous sounds, mainly the following consonant: vowels are longer before lenis than before fortis consonants; before voiced than before voiceless consonants; before fricative than before plosive consonants; before nasal stops than before oral stops, etc. These conditioners of vowel duration seem to operate somewhat similarly in all languages, but further investigation will be made.

Table 7 shows typical variations of vowel duration under all influences that have been investigated in English. Only one of those factors is learned: abridging/expanding (left, vertically); all the others are physiological conditioners: close vowel/open vowel, monophthong/diphthong, voiced/voiceless, stop/fricative, liquid/solid, and oral/nasal. These factors combine to vary the vowel length in a ratio of 1 to 4. This ratio would appear, for instance, in the vowels of: fit (10 cs), five (40 cs).

Comparable tables will be made for the other languages.

7. THE NEUTRAL VOWEL POSITION

This is in principle the vowel one utters when hesitating, when looking for the right word or the next thought. It can also be the average vowel of the weakest stressed syllables, or the implicit vowel toward which all vowels move under unstress. If it is different in each of our four languages, it should be indicative of the center, or basis, of articulation. X-rays and spectrograms are being studied for this sort of comparison.

At present, the point of articulation of the neutral vowel appears to be most fronted in Spanish and most backed in English; its lip rounding most pronounced in French and least pronounced in Spanish; its degree of opening greatest in English and smallest in Spanish. We might tentatively say that the English neutral vowel is somewhat back, lower than mid, and rather unrounded. The German neutral vowel is fairly fronted, mid-open, and between rounded and spread, a little closer to the former. The French neutral vowel is fairly fronted, somewhat higher than mid-open, and rounded. The Spanish neutral vowel is fairly fronted, fairly spread, and between mid-open and mid-close.

It is important to teach this key position at an early stage in the conditioning period.

8. LOSS OF VOWEL COLOR

It seems that in all languages, vowels fail to retain the same timbre when the stress is weak as when it is strong. As a first step toward the comparison of the influence of stress upon vowel color, we are studying by spectrographic analysis the corresponding vowels of such pairs as: abolish/abolition, expecting/expectation, exhibit/exhibition, in substance/insubstantial; Möbel/möbliert, fatal/Fatalismus, Musiker/Musik; ocupan/ocupaban, agradan/agradaban, achican/achicaban; fatigue/fatigant, étude/étudie, provoque/provoquer.

We are interested in finding a) how much change takes place in terms of formant frequency, b) what is the direction of the change -- front, back, center, periphery -- in each one of

the four languages. The articulatory aspect of this change of timbre will also be investigated on motion-picture X-rays.

9. CONSONANT ANTICIPATION

The articulation of a vowel is affected by the anticipation of the following consonant in all languages, but to a very different degree depending on the language. The behavior in the anticipation of post-vocalic consonants and its effects on vowel color, vowel length and syllabic aspect are being studied by cineradiography and spectrography for the comparison of English to Spanish, German and French. The articulatory organs that best permit the observation of anticipation are the lips, the jaws, the tongue tip and the velum. For anticipation of laterals, for instance, the tongue tip is being observed in comparative sentences such as, for English and German: Louis calculated to film the float in the canal/Ludwig kalkulierte, die Flotte im Kanal zu filmen. Distortion of vowel color by consonant anticipation appears to be much more pronounced in English than in the other languages under study.

10. DIPHTHONGIZATION

For the sake of convenience, let us call diphthongs the vowels in which a particular change of color is necessary for comprehension: American English fire [fair], shout [ʃaut], joy [dʒɔɪ] would generally be understood as far [fa:r], shot [ʃa:t], jaw [dʒɔ] without the change of color. And let us call diphthongized vowels the vowels in which no particular change of color is indispensable for comprehension: American English bait, boat, bat, bit, best, boot are not much more likely to be confused with other words when they are pronounced [be:t], [bo:t], [te:t], [bi:t], [bi:t], [bu:t] than [beɪt], [boʊt], [beɪt], [beɪt], [beɪt], [beɪt].

English possesses diphthongs and diphthongizes its monophthongs. German and Spanish possess diphthongs but do not noticeably diphthongize their monophthongs. Modern French does not

possess diphthongs, nor does it diphthongize its monophthongs. Let us treat diphthongs and diphthongization separately.

A. Diphthongs. English, Spanish and German have three diphthongs of the same type: [ai], [au], and [oi], in very broad transcription. This coincidence offers an opportunity for comparing diphthongs among the three languages by cineradiography and spectrography. Comparative sentences of the following type are being used: It's noise in my finest house/In Heuss ist mein feines Haus. Diphthongs seem to differ in three aspects: a) in the color of initial and final target vowels; b) in the relative amount of steady state; c) in the decrease of intensity from beginning to end.

B. Diphthongized vowels. For the acoustic study of diphthongization, a thorough knowledge of spectrography is necessary in order not to confuse formant changes that relate to the perception of diphthongization with formant changes (formant "transitions") that relate to the perception of adjacent consonants. Spectrographic synthesis provides a dependable protection against such causes of error. The use of an electronic time-gate also permits ascertaining which portion of a syllable belongs to the vowel and which to the consonant (or consonants) by perceptual test of separated portions.

For the articulatory study of diphthongization, cineradiography synchronized with sound is an excellent tool. However, a number of subjects must be filmed to find what feature is common to most speakers of a given language. Comparative sentences of the following types are being used: for English and Spanish /e/: It's a day still to face/Que le dé dos cafés; for English and German /e/: In vain they work every day/Für wen gibt es keinen Tee; for English and French /e/: The day before pay wasn't gay/Céder à l'épée n'est pas gai.

A detailed acoustic and articulatory comparison of the vowel types [o], [e], [u], [i] in English and French was completed in our research laboratory and published in The French Review (Vol. 37, pp. 64-76, Oct. 1963). Its figures present the comparison of those four vowels under five aspects: a) the closest phonetic

transcription of the articulatory sequences of the English/French contrasts: know/nos; Pay/(le)ver; do/doux; bee/(ha)bit; b) the profile view sequences of the lips taken at 24 frames per second; c) the front view sequences of the lips at 24 frames per second; d) the X-ray profile sequences of the tongue at 24 frames per second; e) the sound tracks of the eight sequences. The closest transcriptions, on the basis of frame by frame analysis, spectrograms, and auditory study turned out to be: for English know [nnnoooooow], for French nos [nnnooooo]; for English Pay [fffcceeeeeej], for French ver [vvveeeeee]; for English do [dddeuuuuuuuw], for French doux [ddduuuuuu]; for English bee [bbb:iiiij], for French bit [bbbiiii].

The general comparison can be summarized under three aspects.

1. The English vowel is articulated in a constant movement; it changes from frame to frame. The French vowel shows sequences of several frames in nearly complete stability.
2. The English vowel reaches its goal indirectly by way of a wide detour. It starts by rapid over-opening of lips, tongue and jaws (the latter moving late in the initial phase), followed by slow closing and finally over-closing. The /o/ is especially interesting to observe. The tongue constriction displaces itself in a semi-circular movement: it starts in the lower pharynx, rises along the back wall of the pharynx and turns forward toward the velum to end in the front portion of it as for English /u/ but with a narrowness of the pharyngeal cavity that corresponds more nearly to English /w/. The French vowel, on the contrary, reaches its goal directly, without detour, and does not overshoot the mark in the last stage.
3. Energy distribution, as seen on the sound track, is also very different. The English vowel is essentially decreasing, its maximum of intensity being regularly near the beginning. The French vowel does not show any noticeable decrease of intensity; in some cases it even shows an increase.

Observing the progression of the tongue constriction along the walls of the palate, velum and pharynx on these X-ray sequences is indeed indicative of the vowel color changes that occur. But

even more interesting and more closely related to the changes of vowel color are the front and back cavity sizes and openings and their comparison among languages.

11. VOWEL COLOR AND SYLLABLE TYPE

The effect of the type of syllable (closed or open) on the color of vowels seems to be different in each language. It is very marked in French where mid-vowels tend to be open when checked and close when free: j'aime [ʒɛm], j'ai [ʒe], respire [ʁɛs], répit [re]. But it is perhaps not negligible in the other languages. Examples like the following will be studied by spectrography and X-ray motion picture.

English: nesting, netting, pested, petted: dating, day

German: helfen, hellen, Wester, Wetter: Wesen, Weh

Spanish: tez, té; pesca, pesa; percha, pera

12. VOWEL ATTACK AND RELEASE

There is some evidence that English tends to begin initial vowels rather sharply and to end final vowels smoothly, whereas in French initial vowels begin smoothly and final vowels end sharply. It is observed, for instance, that when played backwards English vowels generally have a smooth onset preceded by an [h] sound. French vowels, conversely, have no h and have a rather sharp onset. We need to verify this by analysis and synthesis, to investigate the same points in Spanish and German, and to find how these sharp and smooth vocalic features are produced in the larynx. The behavior of the larynx presents puzzling facts, such as the following. What is heard as a glottal stop attack of initial vowels in German shows two very different types of realization on spectrograms, one by vertical click, indicating a burst or explosion as for stop consonants in general, and the other by a sharp rise in fundamental frequency and no explosion at all.

Since proper attack and release of vowels can play an important role in the auditory impression made by a language,

it is of interest to language teaching and deserves further investigation.

13. VOWEL NASALITY

Vowel nasality can be clearly distinctive, as in French, clearly non-distinctive, as in German and Spanish, and have a doubtful status, as in English.

A. Distinctive nasality. In French nasality is distinctive, it plays a part in changing meaning: vin is different from vais, sombre from sobre, l'un from le, banc from bas, largely because of the nasality in vin, sombre, l'un, banc. (We say "largely" because other factors also contribute to these contrasts -- the tongue position is not quite the same for the orals as for the corresponding nasals; for instance, the vowel of banc is less fronted and less open than that of bas -- see the vowel triangles of Figures 4 and 6.)

In American English, nasality is not generally classed as being distinctive: whether the vowel of name is pronounced [neim] or [nēim] does not affect the meaning. However, in the sequence "vowel + nasal consonant + oral consonant," as in bent [bent], bunt [bant], pond [pand], want [want], haunt [hont], won't [wont], bank [bæŋk], nasality certainly plays a distinctive role; it contributes to the distinction of bent from bet, bank from back perhaps as much as the nasal consonant does. One even wonders, with certain American speakers, whether the nasality of the vowel does not play the major role. Three experiments have been carried out to verify this possibility.

a) We have cut out the [ŋ] of [bæŋk], recorded on magnetic tape by a Midwesterner, and the amputated word sounded more like bank [bæŋk] than like back [bæk]. This indicates that nasality of the æ might be distinctive.

b) We have done a similar amputation on bank and bent by synthesis on artificial speech, and with the same result.

c) We have added artificial nasality by synthesis to the words back and bet, after which naive listeners identified them

as bank and bent, not as back and bet. (Nasalizing a vowel by synthesis is now a standard operation: it consists of decreasing the intensity of the first formant and adding a tone near 250 cps for a fundamental of 120 cps.)

B. Non-distinctive nasality. Nasality being distinctive in French, it is essential that non-nasal vowels show no trace of nasality. In the three other languages nasality being non-distinctive or less distinctive, one does not make a special effort to produce purely oral vowels; the result is that vowels may often be nasalized. We are comparing a) the extent of non-distinctive nasality, b) the place of non-distinctive nasality (generally more before nasal consonants, but elsewhere too) among the four languages. This can be observed on spectrograms by loss of intensity in the first formant, as well as on X-rays by the lowering of the velum.

Sentences of the following types, pronounced by native speakers, are being used to study the nasalization of vowels by anticipation of a nasal consonant. For English and German: The lame lamb is mine/Mein Lamm liegt im Lehm; for English and French: Sam in the cab saw that Ann was sad/On blâme l'Arabe de ce que l'âne est malade and At the camp, the pond sank/La bande campe jusqu'au cinq; for English and Spanish: A fin is on a fish/Afines son las islas. Preliminary results indicate that vowel nasality by anticipation of a nasal consonant is considerable in English, occasional in Spanish, and almost nil in German and French.

In the next and final article, the consonantal features of English, German, Spanish and French will be compared.

COMPARING THE CONSONANTAL FEATURES OF ENGLISH, GERMAN, SPANISH AND FRENCH¹

This is the concluding article in a series of four on the general phonetic characteristics of American English as compared with those of German, Spanish and French. The first three described techniques of research, prosodic features and vocalic features. We are concerned with all aspects of phonetic description that may be helpful to an American teacher in conditioning his students to the study of German, Spanish or French. Our technique of research avails itself of the most advanced development of electronics, such as the manipulation of artificial speech. It may be called four-fold, as it uses cineradiography, spectrographic analysis, spectrographic synthesis, and distributional statistics. Our comparative study of consonants is in progress under the following subdivisions: 1. Articulatory description. 2. Acoustic description. 3. New consonant sounds. 4. (Positional and allophonic) Distribution. 5. Frequency of occurrence. 6. Duration system. 7. Neutral position. 8. Consonant weakening. 9. Vowel anticipation of consonants. 10. Speed of articulation. 11. Tongue fronting. 12. Aspiration. 13. Affrication. 14. Palatalization. 15. Final release. 16. Consonant voicing. Present results are only partial, but it is hoped that an understanding of the direction our research is taking on all these features will benefit the language teacher.

1. THE ARTICULATORY DESCRIPTION OF CONSONANTS

As was the case with vowels, the articulatory description of consonants cannot rely on X-ray still photography. There are two reasons for this. First, no subject can hold a tongue or lip position that really corresponds to a position that is actually

1. The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

met in the course of a consonantal articulation. Second, a consonant is not satisfactorily described by one articulatory position any more than a diphthong can be described by a single position of the tongue; for a consonant is not a stationary position, it is a movement. The outstanding contribution of acoustic phonetic research in the last two decades has been to show that consonants, unlike vowels, are largely perceived by rapid frequency changes in the formants, which reflect rapid changes in the shape and volume of the vocal tract cavities, which in turn are produced by articulatory movements of the tongue, lips, jaws, etc. Experiments with artificial speech demonstrating the transformation of [ga] into [ja], then of [ja] into [i-a] by the mere modification of the rate of change in the formant frequencies are highly revealing of the phonetic nature of consonants. Not only does their production involve movements which are relevant to their perception, but noises of friction [sa, ʃa] or explosion [ba, da], by which the closer consonants have commonly been known, and the articulatory positions of maximal closure by which they have generally been described, seldom play the main role in their linguistic perception. For instance, the burst noise of [ba], the friction noise of [ʃa] do not serve to identify them; it is only the changes in resonance frequency that occur during the opening or closing movements of articulation which are efficient. In short, both place of articulation and manner of articulation of consonants are perceived by means of motional factors which must be taken into consideration in describing them.

Ideal conditions of research will be based on the following precepts.

A. Since consonant articulations are mobile, and not stationary, it is by motion picture, not by still photography, that they must be studied: normal motion pictures for the lips; X-ray motion pictures for the tongue, the palate, the velum, the pharynx, and the vocal tract cavities in general.

B. In order to discover objectively the articulatory move-

ments that are most characteristic of a given language for each distinctive consonant, composite X-ray sequences will be obtained, not by studying a single subject, but by comparing several subjects and finding what they have in common.

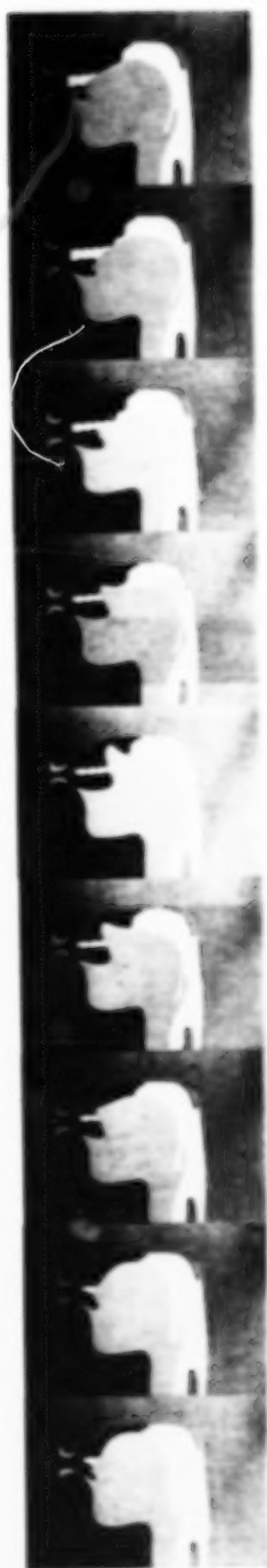
C. To compare the corresponding consonants of various languages, words or utterances in which these consonants are placed in similar phonetic environment must be used as much as is practical. For the comparison of English and French dentals, for instance, motion picture X-rays of the following sentences, spoken by natives, are used: Take the same native shale from the lake in the dale. Les aînés ont aidé chez les Leger l'été passé. To compare English and German dentals: It's deep to his knee. Ein Dieb war er nie. To compare English and Spanish dentals: Notify Tony and Lois Taylor to save and donate a day late or so. Todos los setenta fines no son deleitosos.

As an example of preliminary results of our motion picture X-ray analysis, Figures 1, 2 and 3 present three series of articulatory sequences of /r/ consonants drawn from an animated cartoon we made to compare the /r/ articulation of English, Spanish, and French-or-German. These composite forms are based on the study of many motion picture X-rays for each language, and represent their most characteristic /r/. The French and German fricative pharyngeal /r/ were found similar enough to warrant their representation by the same sequence.

For the prevocalic portion of the American /r/ the mass of the tongue withdraws slightly toward the lower pharynx while the tongue tip rises vertically toward mid-palate, thus creating and enlarging an additional front cavity. In the postvocalic portion the tongue goes to the next vowel by the same route.

The Franco-German /r/ begins with a sharp low and backward movement of the tongue root toward the lower pharyngeal wall, often accompanied by a retraction of the lower jaw. Then the radico-pharyngeal constriction of the tongue rises along the wall of the pharynx before the tongue moves to the position of the next vowel.

Figure 1



English



Franco-German



Spanish

Figure 2



English

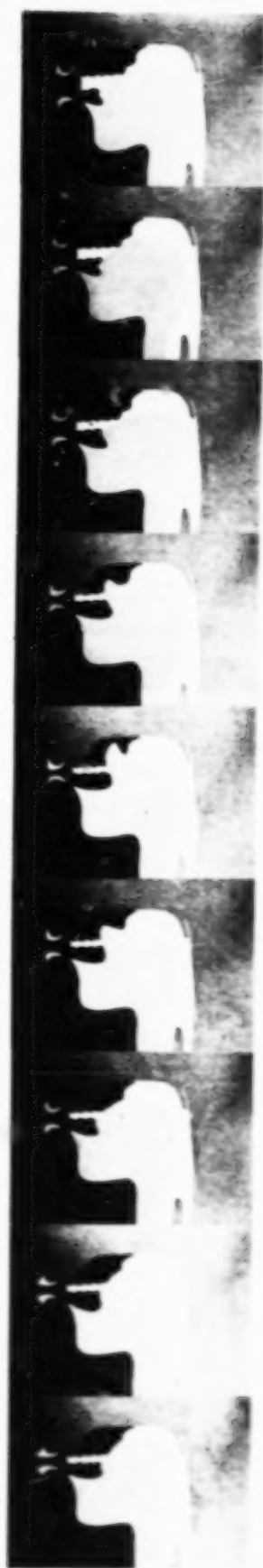


Franco-German



Spanish

Figure 3



English



Franco-German



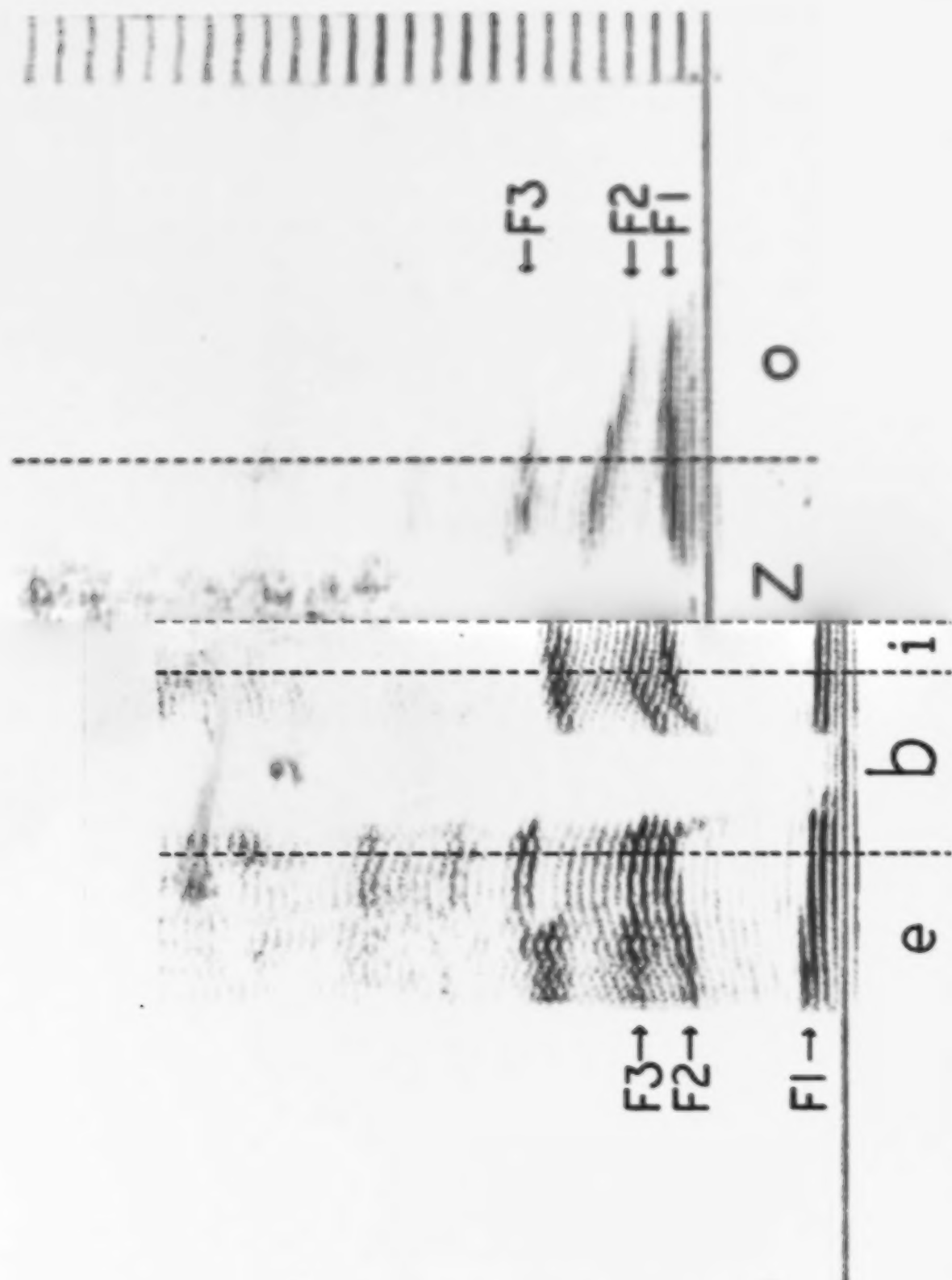
Spanish

For the Spanish /r/ the mass of the tongue remains low and central while the tip moves forward and rises toward the alveols, which it touches just long enough to cause an interruption of the vibrating air flow before moving toward the position of the next vowel.

2. THE ACOUSTIC DESCRIPTION OF CONSONANTS

The fact that consonants are perceived by frequency changes which reflect articulatory movements can be shown with the help of acoustic displays called "spectrograms" even better than with X-rays. Let us examine on Figure 4 the spectrogram of the sounds [ebizo]. ~~A line above the spectrogram indicates the overall intensity.~~ Recent experiments make absolutely clear that the vowels [e, i, o] are perceived only by the short time of near steady-state in the three formants marked F1, F2, F3. (Formants reflect the vocal tract resonances. On the spectrogram of Figure 4 single lines are harmonics; formants are frequency regions of higher energy -- darker regions -- which are formed here of two or three harmonics.) The consonants [b] and [z] occupy all the rest of the spectrogram. The [b] sound is perceived by the falling frequencies of the three formants, corresponding to the closing of the organs; by the near silence that follows, corresponding to the short closure of the organs; by the subsequent burst of noise, corresponding to the brusque reopening of the lips behind which pressure had built up; and by the following rise in frequency of the three formants, which corresponds to the rapid opening of the mouth moving toward the [i] sound. Similarly, the perception of the [z] sound does not depend on the closure portion during which the [z] noise is produced ^{as much as} ~~but~~ on the preceding and following frequency changes in the formants which reflect the changes in cavity shapes and volumes while the mouth is closing and reopening.

It is largely in these formant transitions, therefore, that we shall have to look for consonantal features which distinguish English from the other languages. For example, the fact that the



Figur

tongue is less fronted in English /s/ than in French /s/ is visible on a spectrogram not only in the friction noise, which is lower in English (it sounds somewhat like a /ʃ/ noise to a French ear), but also in the transitions of the second and third formants, which move higher and lower respectively (see Figure 5).

The manipulation of artificial speech provides a most objective means of determining the existence and the extent of differences between two languages. A series of small changes in the formant frequencies and/or in the noises can be made and presented for judgment by ear to two groups of naive listeners from two different countries. While testing a series of small changes going from [asa] to [aʃa], for instance, it was found that some intermediary stimuli were identified as /s/ by American subjects and as /ʃ/ by French subjects. The shift from /s/ to /ʃ/ did not take place at the same frequency for American speakers, whose tongue is less fronted, as for French speakers. In other words, it

discovered and measured by auditory tests. We shall apply this research technique whenever it is feasible and appropriate.

Here, as an example of the use of spectrography for the comparison of consonants, let us content ourselves with comparing the /r/ consonants of the four languages. Figure 6 shows that the main acoustic characteristic of American /r/, the one that distinguishes it from the /r/ consonants of all other languages, is the fall and rise of the third formant, which is responsible for what is well known in linguistic terminology as "r coloring." It corresponds to the creation of a large extra cavity in front by retroflexion of the tongue tip, as shown on Figures 1, 2, 3. For French, German and Spanish /r/ consonants, the third formant remains high and the tongue tip remains low, free of all retroflexion.

Figure 6 also shows that the main acoustic characteristic of French and German /r/ sounds is the rise and fall of the first formant, which indicates that in a resonating tube such as is formed by the vocal tract the maximal constriction is near the

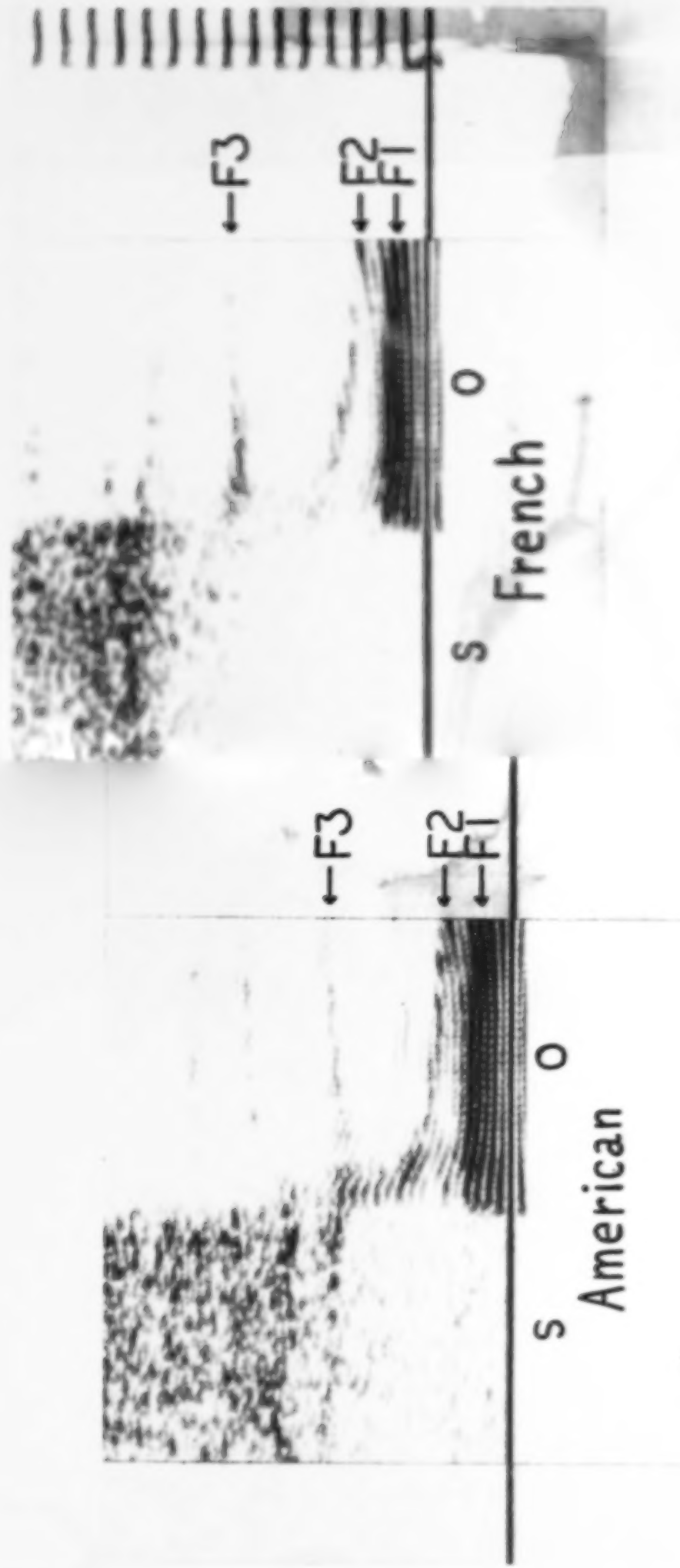


Figure 5

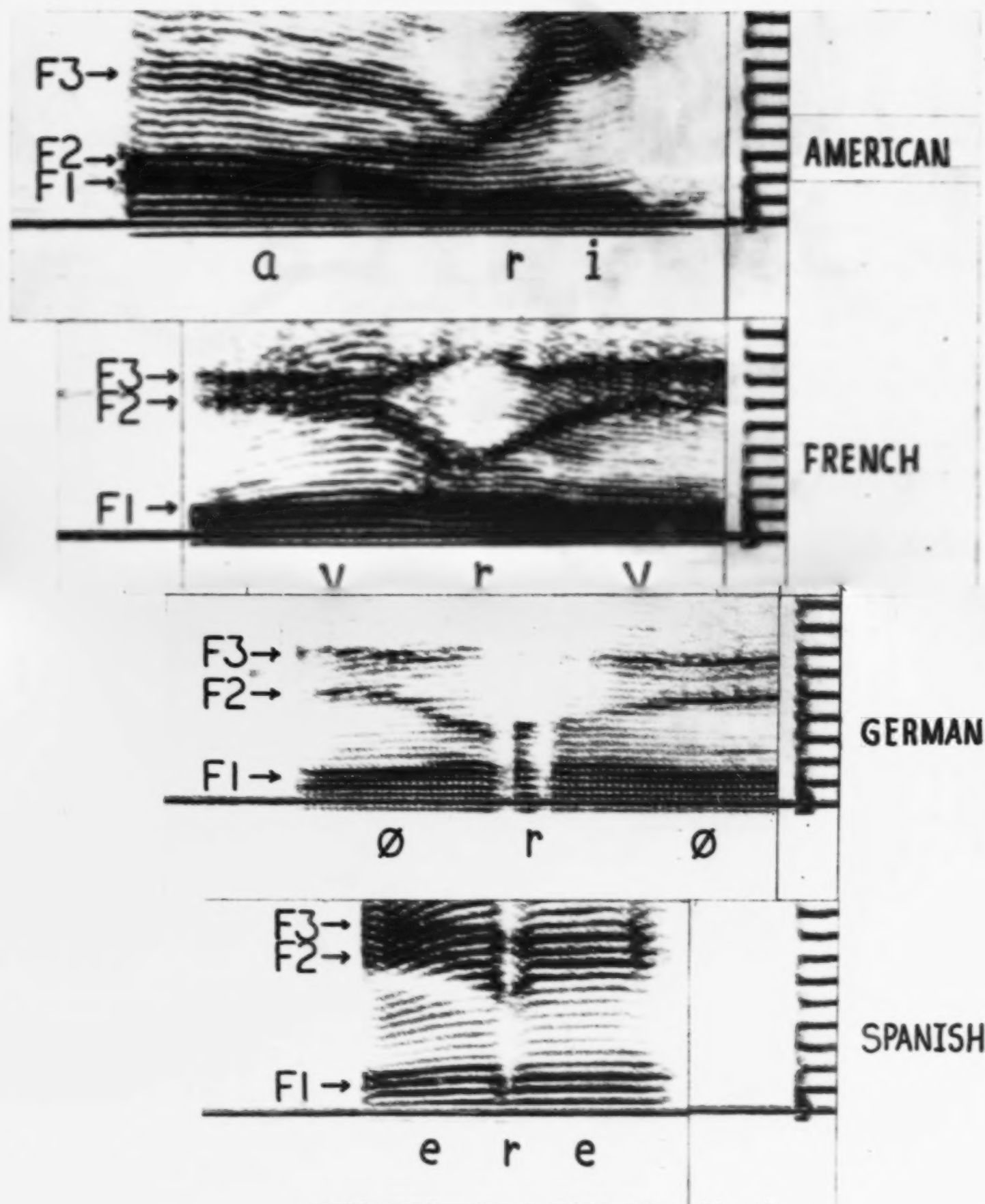


Figure 6

source of sound (the source is at the vocal cords), therefore in the pharynx, as shown on Figures 1, 2 and 3. Besides, the second formant is rather low, an indication that the front cavity is large.

As for the Spanish /r/, it is characterized by a short interruption of vibrations in all formants, corresponding to a short contact of the tongue tip with the alveols, as shown on Figures 1, 2 and 3. This interruption is heard as a flap. Otherwise, the three formants of Spanish /r/ have medium frequencies.

3. NEW CONSONANT SOUNDS

We shall consider here the consonant phonemes of German, French and Spanish which are radically different from those of American English.

Let us first examine Table 1, which offers a complete inventory-comparison among them. Brackets are used when a phonetic realization (allophone) is sufficiently different from the phonemic norm. (The detailed presentation of consonant distribution is in the next chapter.) The letters f and l mean fortis and lenis. These terms describe the p/b type of contrast more universally than do the terms voiceless/voiced -- in Germanic languages, for instance, this contrast may depend more on the presence or absence of aspiration than on the presence or absence of vocal cord vibration, but in either case /p/ is fortis and /b/ lenis.

English is the richest of the four languages with 24 distinctive consonant sounds. None of the other three languages has four distinctive pairs of fricatives -- German and French have only three distinctive pairs, and Spanish has none, unless we consider the positional variants [x], [χ] as contrasting with /θ/, /f/ within the class of fricatives: ador/azor, abanar/afanar. None of the other three languages has a pair of alveolar affricates.

It is only in the category of plosives that the four languages are similar. In the affricates, English and German have two

Table I

Consonant Phoneme Inventory

					English	German	French	Spanish
Stop	Oral	Non-affricate	Labial	{f. l.	p b	p b	p b	p b
			Dento-alveolar	{f. l.	t d	t d	t d	t d
			Palato-velar	{f. l.	k g	k g	k g	k g
		Affricate	Labial			pf		
			Dental			ts		
			Alveolar	{f. l.	tʃ dʒ			tʃ
		Nasal	Labial		m	m	m	m
			Dento-alveolar		n	n	n	n
			Palatal			ɲ		ɲ
	Palato-velar			ŋ			n[ɔ]	
	Bi-labial						b[ɸ]	
	Labio-dental		{f. l.	f v	f v	f v	f	
	Constrictive	Fricative	Dental	{f. l.	θ ð			θ d[ɰ]
			Dento-alveolar	{f. l.	z ʒ	z ʒ	z	s[z]
			Alveo-palatal	{f. l.	ʃ ʒ	ʃ ʒ	ʃ ʒ	
Palatal					x[ç]			
Palato-velar							g[ɣ]	
Pharyngeal					x		x	
Glottal				h	h			
Centro-alveolar							r	
Centro-palatal				r				
Centro-pharyngeal					r	r		
Latero-dental					l	l	l	
Latero-alveolar				l				
Latero-palatal							ʎ	
Palato-spread				j	j	j	j	
Palato-rounded						ɥ		
Resonant	Velo-rounded		w	w	w	w		
	Number of Distinctive Sounds:					24	22	20

phonemes, Spanish one, and French none. (Affricates are considered here as a single phoneme only if they occur initially in native words, and if their friction uses the same point of articulation as their explosion.) In the nasals, the number is the same but not the place of articulation: the two Germanic languages share a velar, and the two Romance languages a palatal. The fricatives vary considerably in number: nine in English, eight in German, six in French, and only four in Spanish, although four more exist as positional variants. In the resonants, English has only one lateral, where Spanish has two; and only two semi-vowels, where French has three.

From the point of view of the number of distinctive consonants, it should be easier for an American to learn French, German or Spanish than for a Frenchman, a German or a Spaniard to learn English!

For French, an American has to learn only three really new consonants: a) The /ɲ/, a nasal stop theoretically articulated with mid-tongue, or dorsum (not the tip -- the tip is suspended toward the lower incisors), in an extended contact with the whole surface of the hard palate, as in gagna. He must avoid raising the tongue tip and making a sequence of n + j. The /ɲ/ is articulated in a single contact and movement using only the mid-tongue (back), not the tip. b) The /y/, a semi-vowel which starts from the /y/ position and abandons it for the next vowel (usually /i/), as in nuît. In /w/, an American is used to combining lip-rounding with tongue-backing. The difficulty here, as for /y/, /ø/, /œ/, is to combine lip-rounding with tongue-fronting. c) The voiced radico-pharyngeal /R/ as in the word mari, normally a very resonant consonant for which a wide constriction is formed between the root of the tongue and the back wall of the pharynx. Occasionally it adds some uvular flaps (hence the name uvular R), or is voiceless with a narrower constriction (hence the classification as a fricative).

For German, an American must learn only four new distinctive consonants: a) The affricates /pf/ as in Pferd and /ts/ as in Zeit. Phonetically, they are true affricates because the friction

that follows the explosion is at the same place of articulation as the explosion -- labial in /pf/, dento-alveolar in /ts/. /ts/ is a new sound for the English speaker in terms of positional distribution. In German it is found in initial position, which is the criterion for an affricate; in English [ts] is found in final position only -- a reason to consider it a cluster. b) The voiceless (fortis) fricative /x/, which has two very different positional variants (see chapter on Distribution): [x] as in lachen, a radico-pharyngeal voiceless friction (between the root of the tongue and the back wall of the pharynx), and [ç] as in sicher, a dorso-palatal voiceless friction (between mid-tongue and the hard palate). Cineradiography will make it possible to find how similar or different are the articulations of German [x], Spanish /x/, and French and German /R/, which all use the pharyngeal wall. c) The voiced radico-pharyngeal /R/, a resonant whose wide constriction is in the same pharyngeal region as that of voiceless [x] or French /R/. It makes frequent use of uvular trill but can also be very open and vocalic (see chapter on Distribution).

For Spanish, an American must learn six new consonant sounds:

- a) Two fricativated positional variants of /b/ and /g/, a voiced bilabial fricative [β] and a voiced dorsal palato-velar fricative [g] as in abogado [aβogaðo]. (See chapter on Distribution.)
- b) A dorso-palatal nasal stop /ɲ/, as in mañana [manana], articulated with mid-tongue (dorsum) in extended contact with the entire hard palate. The tip of the tongue is down behind the lower teeth and does not rise as it does in [nj] of onion in English.
- c) A dorso-palatal lateral resonant /ʎ/ as in calle [kaʎe], articulated with the center of mid-tongue in extended contact with the center of the hard palate. The tip of the tongue is down behind the lower teeth and does not rise as it does in [lj] of will you in English.
- d) A voiceless radico-pharyngeal fricative /x/, normally resulting from a narrow constriction between the root of the tongue and the back wall of the pharynx.
- e) A voiced apico-alveolar (tongue tip) flapped /r/, a resonant with one flap when it is single, or short, as in pero, and multiple flaps when it is double, or long, or geminated, as in perro.

Summary of new consonant sounds
for an American student

In French:	p, ɥ, R	(3)
In German:	pf, ts, x[x], x[ç], R	(5)
In Spanish:	p, ʎ, b[β], g[ɣ], x, r	(6)

For German, as well as for French and Spanish, an American student finds new sounds in the pharynx: the voiced French and German [R] sounds, the voiceless German and Spanish [x] sounds. As there exist no pharyngeal articulations in English, the American student has to acquire the new articulatory habit of tongue root withdrawal toward the lower region of the pharyngeal wall. This must be done with tongue tip low if the frontal resonance which characterizes Spanish as well as German and French is to be produced. Both tongue root backing and tongue tip lowering are opposed to the American habit of centering the mass of the tongue and raising the tip in a retroflex motion, which produce the characteristic r-coloring of American English.

4. CONSONANT DISTRIBUTION

To compare the distribution of English consonants to that of German, French and Spanish consonants, it is convenient to consider separately the positional distribution of phonemes and that of allophones. We shall limit ourselves here to single consonants.

A. Phonemic distribution of consonants

Tables 2, 3, 4 and 5 present, in examples, the positions in which single consonant phonemes of the four languages occur.

English does not have initial /ɔ/ and /ʒ/, or final /n/, /w/, /j/. German does not have initial /s/, /x/, /ɔ/, or final /b/, /d/, /g/, /v/, /z/, /j/. French does not have initial /p/, or final /w/, /ɥ/. Spanish is very weak in final consonants and especially

English Consonant Inventory

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
p	pin	mopping	tap
b	bin	mobbing	tab
t	tin	debtor	cat
d	din	deader	cad
k	kale	vicar	back
g	gale	vigor	bag
tʃ	chin	breeches	batch
dʒ	gin	bridges	badge
m	mail	simmer	tam
n	nail	sinner	tan
ŋ		singer	tang
f	fail	safer	half
v	vale	saver	have
θ	thigh	ether	loath
ð	thy	either	loathe
s	seal	ricer	bus
z	zeal	riser	buzz
ʃ	shale	glacier	cash
ʒ		glazier	rouge
h	hail	ahead	
r	rain	around	care
l	lane	alone	tall
j	yacht	beyond	
w	watt	away	

Table 2

Table 3
German Consonant Inventory

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
p	Panne	Oper	knapp
b	Bann	Ober	
t	Tanne	Otter	Schutt
d	dann	oder	
k	Kanne	Hacken	Rock
g	ganz	hagen	
pf	Pfanne	Opfer	Knopf
ts	zanken	setzen	Putz
m	lagen	Eimer	lahm
n	nagen	einer	Lahn
ŋ		hängen	lang
f	Fahne	schlafen	auf
v	Wanne	Löwe	
s		wissen	Schuss
z	Sahne	lesen	
ʃ	Schande	Tasche	Tausch
ʒ	Genie	Gage	Rouge
ç	Chemie	sicher	reich
x		lachen	Rauch
h	Haus	Vorhang	
r	Rand	Säure	vier
l	Land	Säule	viel
j	jagen	Boje	

French Consonant Inventory

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
p	paix	apport	type
b	baie	abord	robe
t	temps	coûter	vite
d	dans	coudre	vide
k	quai	écu	roc
g	gai	aigu	rogue
m	mes	renis	femme
n	naïf	renie	fane
ʒ		oignon	peigne
f	fait	raffine	neuf
v	vais	ravine	neuve
s	celle	poisson	race
z	zèle	poison	rase
ʃ	chez	boucher	franche
ʒ	jet	bouger	frange
r	rez	courant	soeur
l	les	colère	seul
j	yeux	ailleurs	fille
ɥ	huit	cruel	
w	oui	ébloui	

Table 4

Spanish Consonant Inventory

	<u>Initial</u>	<u>Medial</u>	<u>Final</u>
p	peso	capa	
b	beso	cava	
t	tía	rata	
d	día	rada	dad
k	casa	vaca	
g	gasa	vaga	
tʃ	chía	racha	
m	mueve	lama	
n	nueve	lana	dan
ɲ		laña	
f	foco	rifa	
θ	zoco	riza	voz
s	sota	casa	das
x	jota	caja	boj
r	rama	pero	mar
l	lama	pelo	mal
ʎ	llama	pello	
j	yeso	cayada	
w	hueso	ahueca	

Table 5

in voiceless ones. The only two final voiceless consonants are /s/, preserved by the plural, and /x/, with very low frequency of occurrence. It has only one final plosive versus six in initial position, and it is fricativated into [x] -- in other words, it has no plosive allophone in final position. It has only one nasal, /n/, in final position, versus three in medial position; only two resonants, /r/, /l/, versus five in initial or medial position. It is richest in final fricatives, /θ/, /s/, /x/, three versus four in initial or medial position.

Summary: The only position in which all consonants occur in each language is the intervocalic (medial). Palatal and velar nasals practically do not occur initially, and the semi-vowels practically do not occur in final position in all four languages. German is characterized by its lack of voiced plosives and voiced fricatives in final position. Spanish has very few final consonants -- only seven out of the nineteen Spanish consonant phonemes occur in final position.

B. Allophonic Distribution of Consonants

1. English Allophonic Distribution.

a) Aspiration. Fortis plosives /p/, /t/, /k/, are aspirated before a vowel in initial and/or stressed syllables: potato [phəθetə], cool [khu], but are unaspirated in all other positions and after [s]: happy [hæpi], specific [spesɪfɪk], school [sku], stale [steɪ]. This does not occur in French and Spanish. An American will have to pronounce unaspirated /p/, /t/, /k/ in all positions: capitale, patata. It does, however, occur in German, where the "after [s]" rule of no-aspiration even extends to [ʃ]: Skat [skath], stehen [ʃteen], and aspiration extends to final consonant release: [skath].

b) The American /l/ is less fronted in post-vocalic, final or preconsonantal position: hell, help, helping (dark l) than in prevocalic, initial or post-consonantal position: lace, place (clear l). This does not occur in any of the three target languages. An American will, therefore, have to learn to pronounce

clear l in post-vocalic positions in German, Spanish and French: helfen, alba, calmé.

Figures 7 and 8 offer a comparison of prevocalic and post-vocalic /l/ in the four languages. They show the position of maximal closure found on motion picture X-rays at 24 frames per second. The releasing /l/ and the arresting /l/ to be compared are taken in each language from words in which they are adjacent to a similar vowel: for English, Louis and Paul (two back-rounded vowels); for German, Lam and Kanal; for Spanish, la and canal; for French, laisse and telle. In each case, the same subject articulates both /l/ sounds. The articulations are those of a single subject in each language; they do not represent the composite form of many subjects, which will eventually emerge when more X-ray films are available. Comparisons of the frames shown in Figures 7 and 8 oppose English to the three target languages. For the postvocalic /l/ of the American subject, the tip of the tongue is more retroflexed, the dorsum is more concave, and the mass is more backed than for the prevocalic /l/. For the German, Spanish and French subjects, differences between prevocalic and postvocalic /l/ articulations appear to be negligible. If anything, the postvocalic /l/ of the German subject is slightly more fronted than the prevocalic one.

c) The Western American /r/ can be articulated in two different ways, known as retroflexed and bunched. They are identified by ear as similar sounds, and both cause a lowering of the third formant on spectrograms. For the retroflexed /r/, the palatal constriction is produced by raising back the tip of the tongue toward the hard palate. For the bunched /r/, the palatal constriction is produced by raising back the mid-dorsum of the tongue toward the hard palate, leaving the tip down or absorbed in the ball-shape of the mass of the tongue. Some American speakers use the retroflexed /r/ in all positions, others use the bunched /r/ in all positions. Still others use both types of /r/, and these speakers use the retroflexed /r/ in strong (prevocalic) position, and the bunched /r/ in weak (postvocalic or

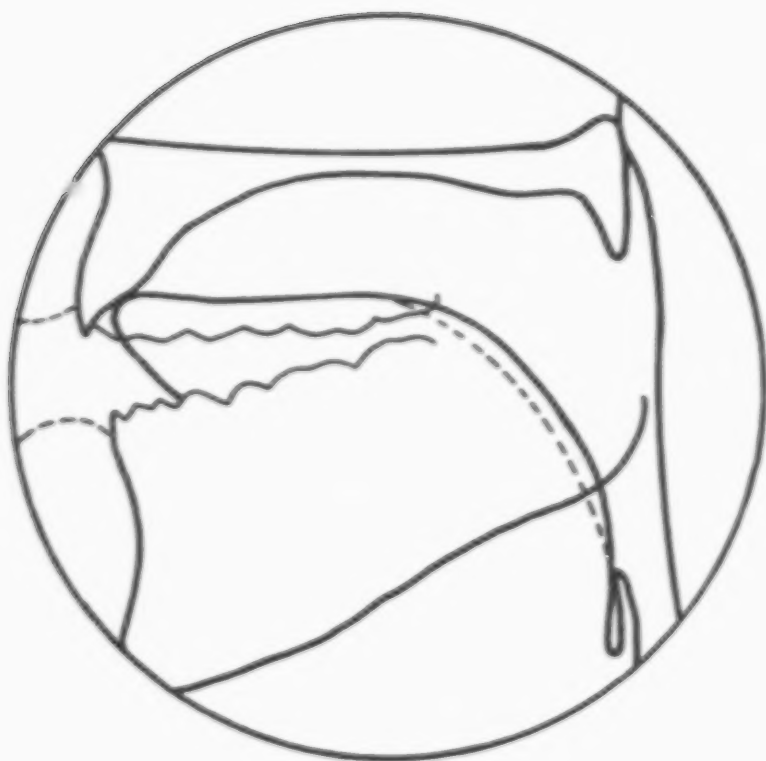
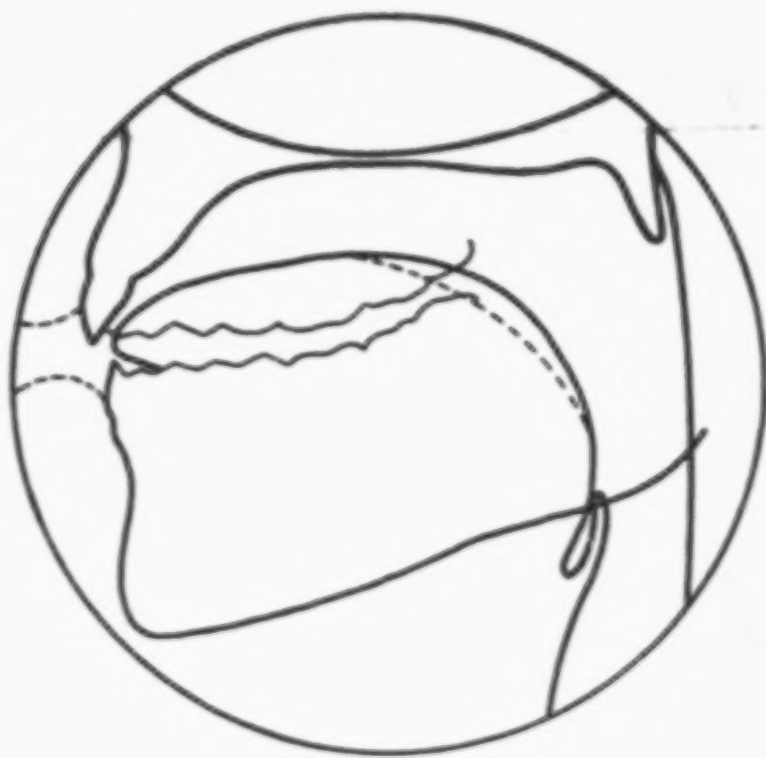


Figure 7

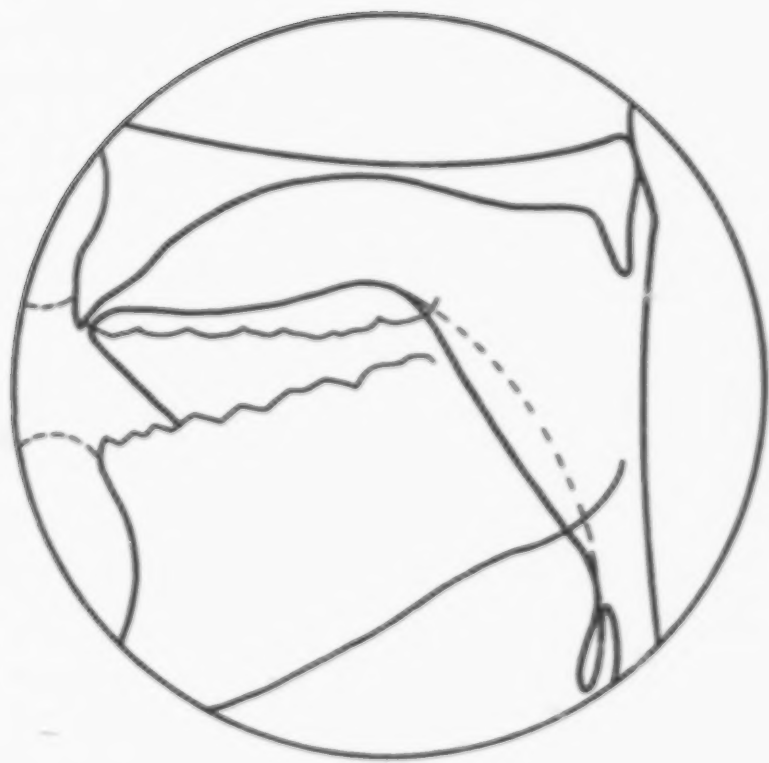
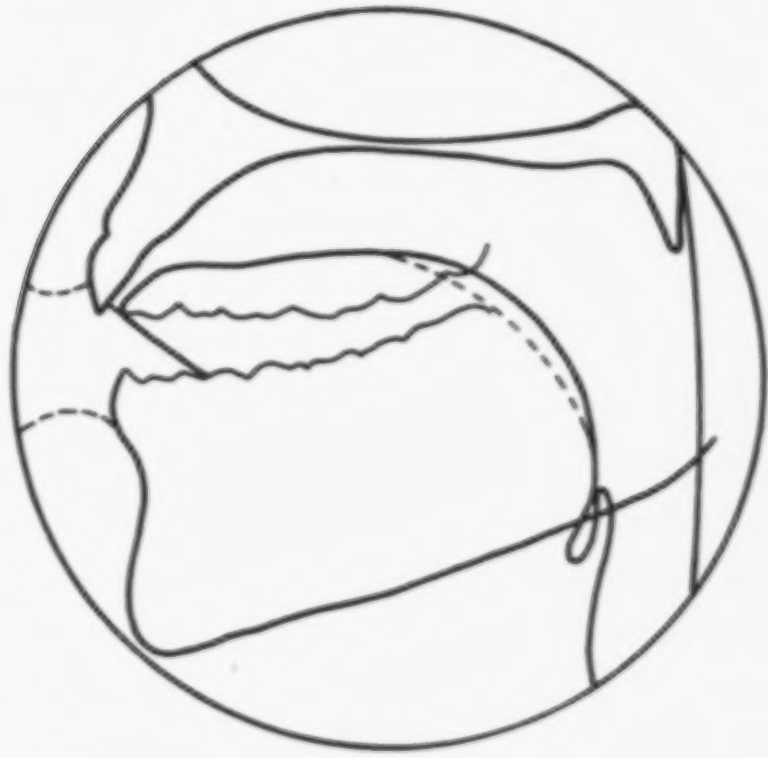


Figure 8

intervocalic) position.

Figure 9 presents for comparison two motion picture sequences, one of retroflexed /r/ in strong position and one of bunched /r/ in weak position. One can see that both articulations shape a pre-tongue cavity and that these pre-tongue cavities are of comparable sizes -- a fact that accounts for comparable lowering of the third formant on spectrograms.

2. German Allophonic Distribution

a) Aspiration. The fortis plosives /p/, /t/, /k/ are aspirated before all vowels and in final release: tot [thoth], but are not aspirated after /ʃ/ and /s/: stehen [ʃteən], Skat [skath]. On one hand, aspiration extends farther than in English; on the other hand, it is prevented much more frequently by a preceding /ʃ/ than by a preceding /s/. (See English aspiration rules in section on English distribution.)

b) The voiceless fricative phoneme /x/ is articulated with a dorso-palatal constriction [ç] after front vowels and after all consonants: Ich [iç], Bücher [byçər], Milch [milç], and with a radico-pharyngeal constriction after back vowels, including the [a] sounds: Nacht, [naxt], Buch [bux], auch [aux].

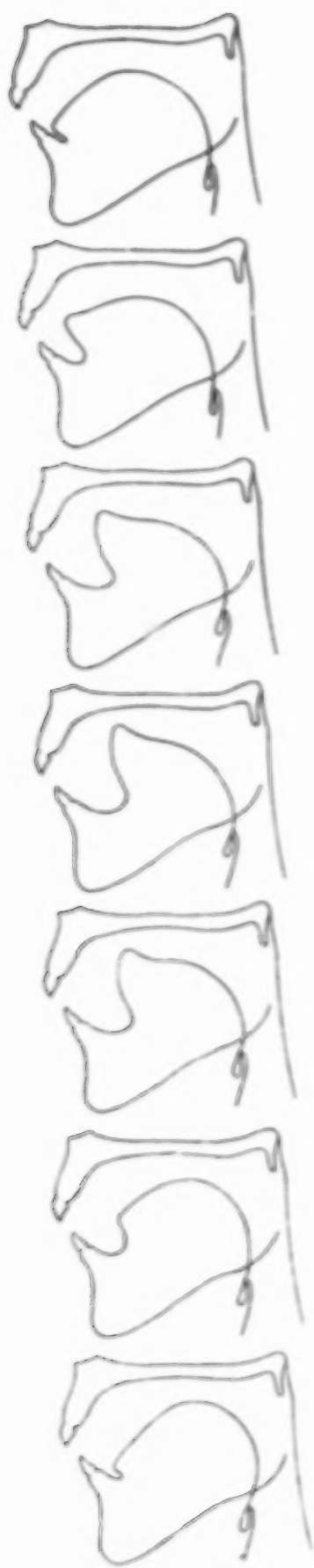
c) All final consonants unvoice: Tag [tak], sind [zint]. The final voiced-voiceless contrast of English bag/back does not occur.

d) All syllable initial s are voiced: sind [zint], so [zo]. The initial s/z contrast of English seal/zeal does not occur.

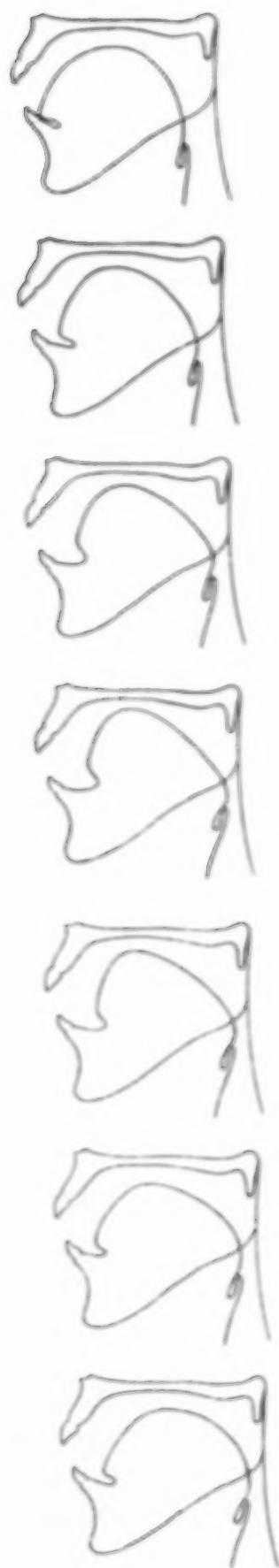
e) After short stressed vowels, consonants are more arresting ^{often} and hold the closure longer: bitte [bɪt:ə] than after long vowels: biere [bi:tə]. In English, consonants that follow short stressed vowels tend, on the contrary, to be very short: bitten [bɪtən].

3. French Allophonic Distribution

a) Latent consonants (suggested in spelling by a word final silent consonant) are sounded with the following initial vowels
1) when a determinative precedes a noun: les enfants, petit enfant; 2) when a personal pronoun precedes or follows a verb:



Retroflexed



Bunched

Figure 9

vous avez, ont-ils; 3) when a monosyllabic preposition or adverb precedes the word it determines: sans elle, très utile.

b) Latent consonants are silent 1) before a singular noun: soldat/anglais; 2) before an aspirate h (which is silent but prevents liaison as well as elision): les/héros; 3) before numerals un, huit, onze: cent/un; 4) after et: et/elle.

c) Latent consonants are silent in all other cases unless the style is formal: Allez ensemble is formal, Allez/ensemble is familiar.

d) Unstressed syllables (non-final of sense groups) ending in a consonant cluster (two consonants or more) and followed by a syllable initial consonant tend to append a schwa to the cluster: ours blanc [uʁsəblɑ̃], vendredi [vɑ̃dʁedi], une fenêtre [ynfənɛtʁ].

4. Spanish Allophonic Distribution

a) The voiced consonants /b/, /d/, /g/ are pronounced as plosives in initial position: bola [bola], after nasals: cuando [kwando], and /d/ after /l/: caldo [kaldo]; but they are pronounced as fricatives in all other positions of the word or phrase: abogado [aβogaðo], pierde [pjerðe]. In English, /b/, /d/, /g/ are stop consonants in all positions.

b) Spanish /s/ tends to voice before voiced consonants: isla [izla], asno [azno]. In English, /s/ can remain voiceless in similar positions: wrestler [rɛslər], misname [mɪsnem].

c) Spanish /n/ takes the "place of articulation" of the following consonant: sin darlo [sin], sin pagar [sim], sin correr [siɲ]. In English, /n/ keeps its proper place of articulation regardless of the following consonant: in part, in care of.

d) The Spanish non-geminated /r/ is pronounced with a single flap in intervocalic position, and with multiple flap in initial position: raro [rraro]. This non-distinctive difference is about the same, phonetically, as the distinctive difference that occurs only in intervocalic position: pero /pero/, perro /perro/. Western American /r/ is nearly the same in all positions: rarer.

5. CONSONANT FREQUENCY OF OCCURRENCE

The auditory impression a foreigner receives of a language when he enters the country in which it is spoken could very profitably be recreated for the student in the initial stage of learning if the teacher were informed about consonant as well as vowel frequency of occurrence. The learning sequences that are used to initiate a student to the articulatory habits of a language should ~~also~~ ^{partly} be based on frequency of occurrence of consonants and vowels. To make this possible, Table 6 offers a comparison of the frequency of occurrence of consonants among the four languages.

These consonants take more place in the auditory impression of German or English (63.7% and 62.6% of all phonemes) than in that of French or Spanish (56.8% and 56.4% of all phonemes). In addition to the occurrence of phonemes, Table 6 shows the proportion of allophones that are markedly different.

On the surface the four languages do not differ strikingly in the frequency ranking of their consonants. In English and German the six highest phonemes are similar: /t, n, r, l, s, d/ in English, /n, t, r, d, s, l/ in German. And the most frequent consonants of French and Spanish are not markedly different -- they include the same apical sounds /t, s, l/. The first four consonants include these three apicals and the /r/ in each language.

But the auditory impressions produced by these consonants will diverge much more than their symbolization shows. The /r/, which is among the three highest consonants in all four languages, is of one type in American English, another in Spanish, still another in French and German. Here French and German are together with a similar pharyngeal constriction of high recurrence.

The /l/, which is among the seven highest consonants in all four languages, will set the auditory impression of English apart from the others.

However, it must be noted that after /r/ and /l/ the other characteristic articulations of each language are rather in the lower half of the table. (The same had been noted for vowels;

Table 6

Comparative Frequency of Occurrence of Consonants

English		French		German		Spanish	
Cons.	%	Cons.	%	Cons.	%	Cons.	%
t	12.77	r	15.00	n	16.76	s	14.60
n	11.46	l	10.62	t	11.49	n	12.27
r	8.32	t	9.67	r	11.10	r	10.88
l	7.69	s	8.76	d	8.24	d	9.08
s	7.47	p	7.96	ʃ	6.80	t	8.30
d	5.65	d	7.23	l	6.08	k	8.11
z	4.90	k	6.35	x	5.32	l	6.72
m	4.74	m	5.98	m	5.23	m	6.53
θ	4.61	n	5.22	f	3.94	j	5.14
k	4.30	v	4.45	v	3.88	b	4.99
w	3.67	j	3.21	g	3.76	p	4.02
b	3.48	ʒ	2.88	z	3.67	θ	2.48
h	3.26	f	2.55	b	2.86	w	2.18
v	3.17	z	2.34	k	2.29	g	1.24
f	2.86	w	2.30	ts	2.26	f	0.90
p	2.35	b	2.26	ʃ	1.93	ʎ	0.83
ɒ	2.20	g	1.13	h	1.90	x	0.64
j	2.01	ʝ	0.99	p	1.02	tʃ	0.56
g	1.57	ɥ	0.84	ɔ	0.81	ɲ	0.53
θ	0.97	ɲ	0.26	j	0.54		100.00
ʃ	0.88		100.00	pf	0.12		
dʒ	0.88				100.00	/b/-[b]	1.27
tʃ	0.63			/x/-[x]	1.62	\"[ɸ]	3.72
ʒ	0.16			\"[ç]	3.70	/d/-[d]	2.96
	100.00					\"[ɖ]	6.12
/l/-[l]	4.08					/g/-[g]	0.30
\"[ɭ]	3.61					\"[ɣ]	0.94
						/s/-[s]	13.29
						\"[z]	1.31
						/r/-[r]	9.95
						\"[r:]	0.93
						Preconsonantal only	
						/m/	0.52 before labials
						/n/-[ɔ]	0.71 before velars
						\"[n]	4.43 elsewhere

the most characteristic ones -- French nasals, German and French front-rounded, Spanish, German and English diphthongs -- all had low frequency of occurrence.)

Of the characteristic consonants of English, /θ/ is still in the upper half, because of the article the; /h/ is midway; /ʊ/, /θ/, /tʃ/ and /dʒ/ are near the bottom.

Of the characteristic consonants of French, /ɜ/, /y/ and /j/ are in the lower half, the last two at the very bottom.

Of the characteristic consonants of German, [ç], [x], /ts/, /pf/, /n/ and /ŋ/ are in the lower half.

Of the characteristic consonants of Spanish, [ɲ] is in the upper half, but the others, [ʎ], /θ/, [g], /k/, /x/, /tʃ/, /j/, [ɔ], are all in the lower half of the ranking by frequency of occurrence.

To summarize, the utility of a frequency of occurrence list such as the one on Table 6 is not so much to show the ranking of consonants ~~that are more or less shared by the four languages~~ -- these do not differ noticeably from language to language -- as to provide a ranking of the characteristic, or unusual, articulations of each language.

In some cases, similar symbols hide very different articulations, as for the /r/ consonants of all four languages, the /l/ of English versus the three other languages, the /x/ of German versus Spanish, the /b/, /d/, /g/ and /n/ of Spanish versus the three other languages.

In creating material that will convey the characteristic auditory impression of a language, the following consonants should be emphasized in decreasing order of importance:

English:

r (8.32), ʃ (4.31), h (3.26), ʊ (2.20), θ (.97), dʒ (.88), tʃ (.63)

French:

R (15), ɜ (2.28), y (.84), j (.26)

German:

R (11.10), ç (3.70), ts (2.26), h (1.90), x (1.62), ɔ (.85), pf (.12).

Spanish:

r (10.00), ʎ (6.12), ʝ (3.72), θ (2.48), ɣ (.94), ʎ (.83),
 ɟ (.71), x (.64), tʃ (.56), ɲ (.53).

As it may sometimes be useful to visualize frequency of occurrence of all phonemes together, regardless of whether they are consonants or vowels, Table 7 was prepared. Allophonic separations are not included. It is certainly indicative of the phonemic composition of the most frequent syllables, ~~possibilities for various syllables~~, but for this purpose it is safer to refer to our earlier treatment of syllable types under the general title of Prosody.

6. DURATION SYSTEM

Whereas vowel duration is very relevant within words in German, and fairly so in English and French, consonant duration plays a distinctive role in only two of our four languages, and it is a minor role.

A. Distinctive consonant length

In intervocalic position only, the duration of Spanish /r/ is distinctive: pero, caro are distinguished from perro, carro by the length of the /r/ consonants. As shown on Figure 10, the short /r/ has only one flap (interruption) and the long /r/ has several flaps -- most frequently three or four.

In intervocalic position only, the duration of French /r/ is distinctive: the imperfects il courait, il mourait, il acquerrait are distinguished from the conditionals il courrait, il mourrait, il acquerrait by the length of the /r/ consonants. Figure 11 shows an average difference of length between short and long /r/ in French. The phenomenon is limited to these three verbs, and verbs ending in vowel + rer, and their compounds.

The dialect of Paris makes another distinction by means of consonant length. In intervocalic position only, the /l/ of the personal pronouns le, la, les is significantly longer than all

Table 7

Comparative Frequency of Occurrence of Phonemes

English		French		German		Spanish	
Vowels	Cons.	Vowels	Cons.	Vowels	Cons.	Vowels	Cons.
a	8.87		8.67		10.45	e	14.00
	7.85	e	8.14	a	8.99	a	12.97
	7.04	a	7.04		7.17	o	9.21
i	5.57		6.14		6.92		8.37
	5.11		5.59		5.14		7.03
	4.72	i	5.23	i	4.33		6.24
	4.59		5.06		4.24		5.21
æ	3.64		4.60	a	4.03		4.75
	3.47		4.18		3.79		4.65
i	3.28		3.67		3.32	i	4.47
	3.01		3.46		3.26		3.85
	2.91	e	3.21	ai	2.96		3.74
	2.83	æ	3.20	ɛ	2.66		2.95
ɑ	2.70		3.02	e	2.53		2.86
	2.64	ɛ	2.83		2.46		2.30
ɛ	2.64	u	2.70		2.42	u	1.81
	2.26		2.57		2.35		1.42
u	2.16	o	2.13		2.29		1.25
	2.14	y	1.98	i	2.21		0.71
ai	2.12		1.86	u	1.99		0.52
	2.01		1.67		1.78		0.47
	1.95	ɜ	1.62	a	1.46		0.37
o	1.91		1.48	u	1.44		0.32
	1.75		1.35		1.43		0.30
æ	1.70		1.33	o	1.41	ei	0.09
e	1.52		1.31		1.41	ai	0.06
	1.45	o	1.10		1.20	au	0.06
	1.35	ɛ	1.03	au	1.18	oi	0.02
	1.23	æ	0.76		1.18		
	0.96	e	0.72	o	0.86		
au	0.84		0.65		0.64		
ɔ	0.77		0.57		0.51		
u	0.77		0.49	y	0.49		
	0.60	æ	0.44	Y	0.38		
	0.54		0.15	e	0.36		
	0.54	a	0.01		0.34		
	0.39			oy	0.21		
	0.10			æ	0.13		
oi	0.08				0.07		
	100.00		100.00		100.00		100.00

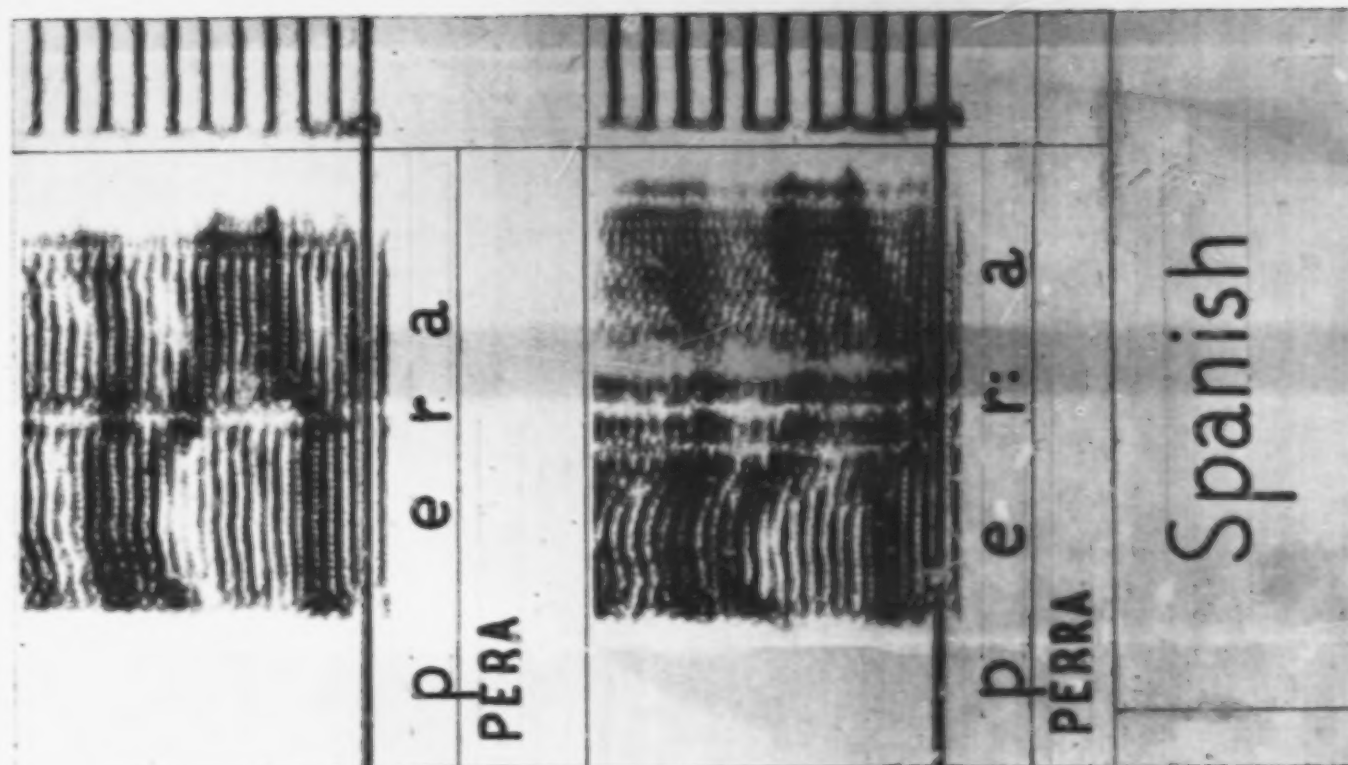


Figure 10

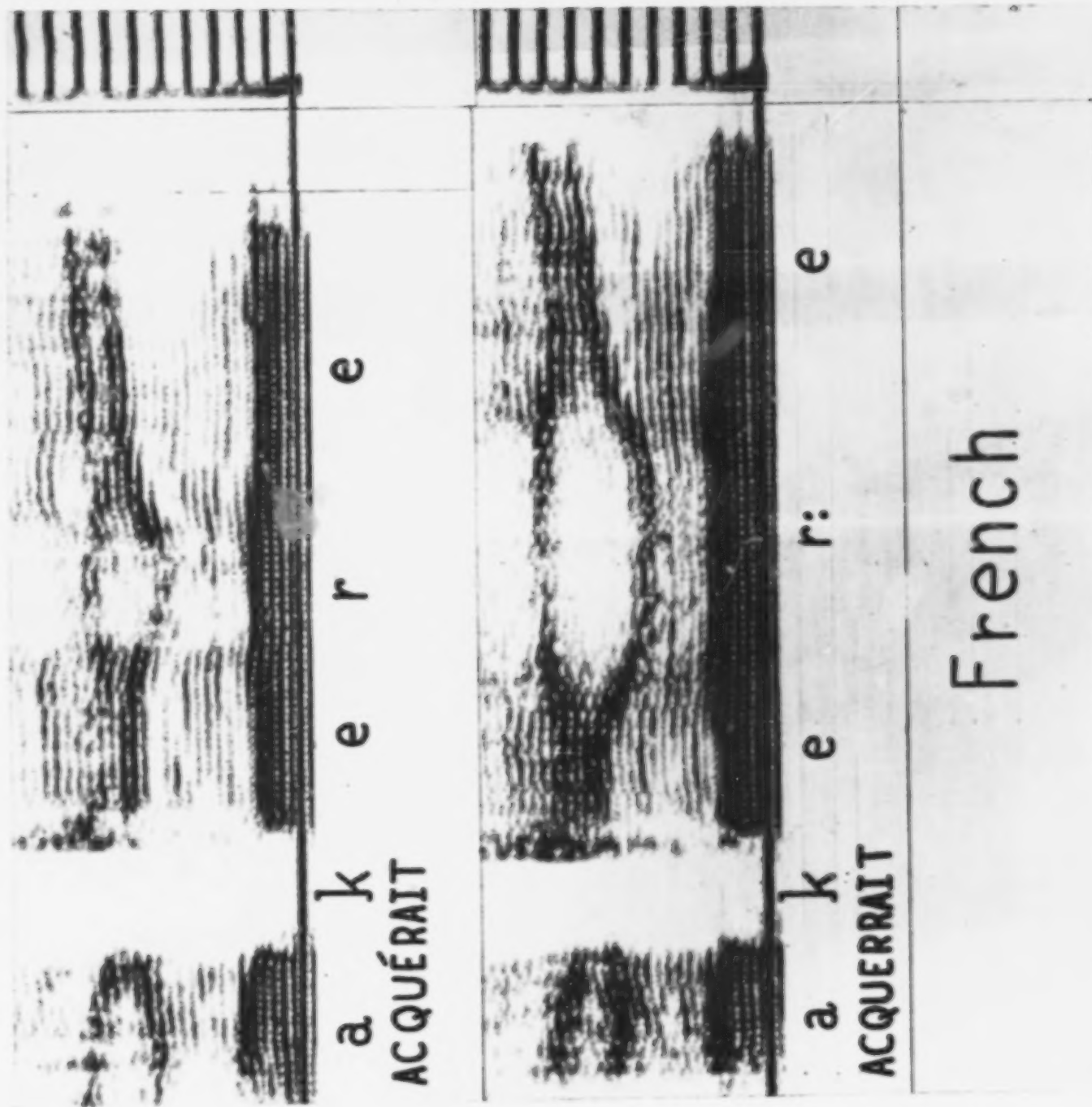


Figure 11

other intervocalic /l/ phonemes. The /l/ is longer in Ma l'a dit [malladi] than in maladie [maladi], in celui qui l'a vu [sɛlɥi-killavy] than in celui qu'il a vu [sɛlɥikilavy]. Figure 12 shows an average difference of length between short /l/ and long /l/, and a comparison with double (geminated) /l/ in le mal l'a dit. The phonetic realization of the geminated consonant shows greater consonant anticipation -- the arresting transitions from /a/ to /l/ are more marked.

In German and English, consonant length is never distinctive in words. Naturally, the gemination of consonants between words is equally distinctive in all four languages: key players [ki-plɛɹz]/keep players [kipplɛɹz]; ton bier [tɔ̃bjɛʁ]/tombe bien [tɔ̃bbjɛʁ]; boca libre [bokalibre]/vocal libre [bokallibre]; neue Nadeln [nɔyənadɛln]/neuen Nadeln [nɔyɛnnadɛln].

B. Non-Distinctive Consonant Length

In the chapter on Vowel Duration (IRAL 2/2), it was shown that in all languages the length of vowels can vary considerably under the influence of a following consonant, and that conditioning of vowel length never is distinctive. It was said that, all other conditions being equal, the length of a vowel depends on the articulatory force of the following consonant, that is, on the amount of energy involved in articulating that consonant -- the stronger the consonant, the shorter the preceding vowel. For instance, within each class, the fortis consonants shorten the preceding vowel more than the lenis ones: the vowel is shorter in bit than in bid, in beat than in bead, in back than in bag, in rich than in ridge, in race than in raise, etc.

But our investigations also show that strong consonants have longer closure than weak ones. So that the duration behavior of consonants is the reverse of that of vowels -- shortened vowels are followed by longer consonant closures, and lengthened vowels are followed by shorter consonant closures. When consonants, therefore, are compared within each class, it is found that, all other conditions being equal, fortis plosives [p t k] have longer closure than lenis plosives [b d g], fortis affricates [tʃ ts] than lenis affricates [dʒ dz], fortis fricatives [f θ s ʃ x] than lenis fricatives [v ʒ z ʒ R].

Figure 12

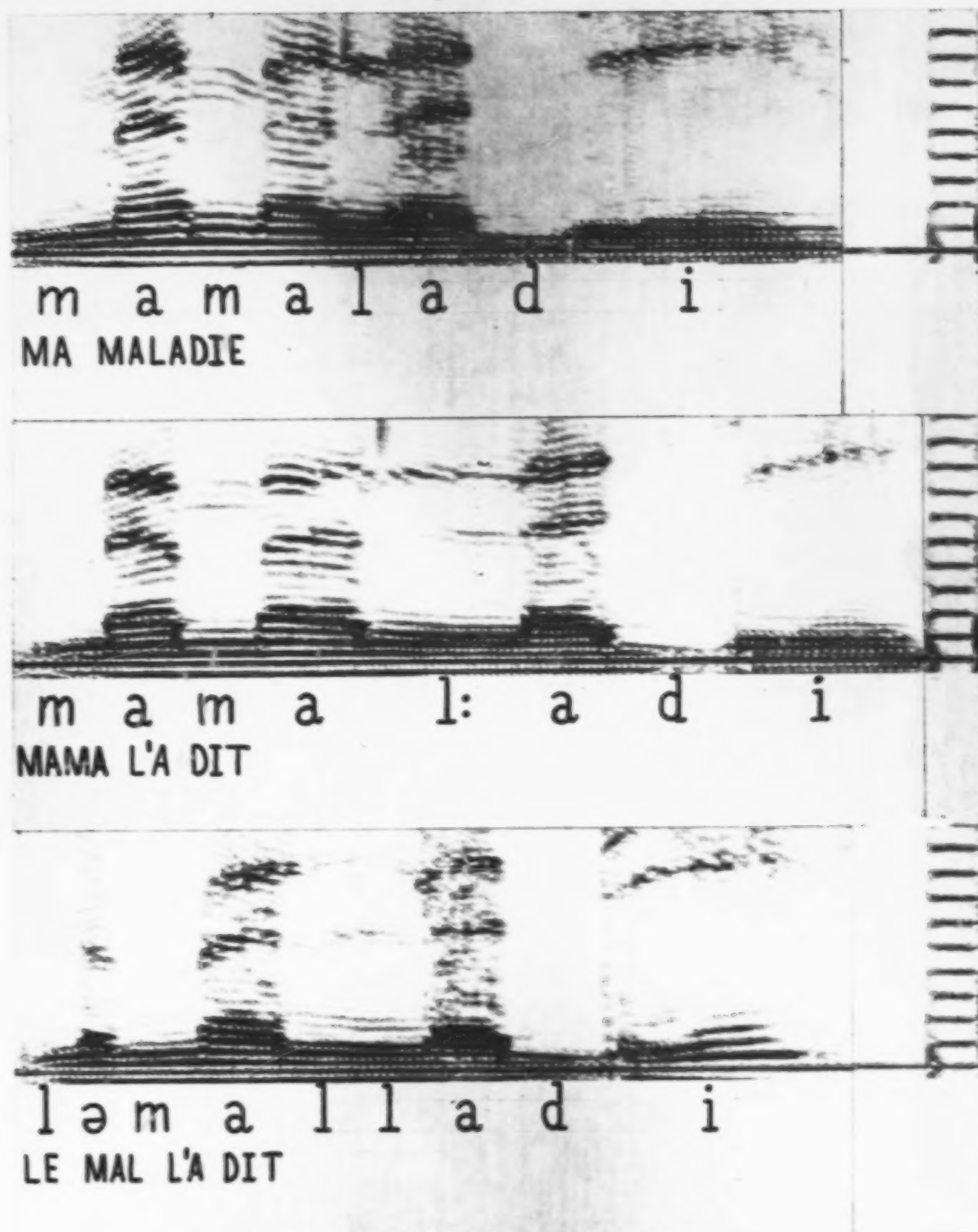


Figure 13 shows that non-distinctive differences in closure of consonants exist in English as well as in German and in French. (The fortis/lenis contrast does not occur in postvocalic position in Spanish.) But they are more pronounced in final than in non-final syllables.

7. THE NEUTRAL CONSONANT POSITION

The most characteristic shape of the tongue for each language is expected to appear through an analysis of cineradiograms of key consonants, and of rest positions to which the tongue tends to return. One of the key consonants is /l/.

A preliminary comparison of tongue shape for /l/ is presented in Figure 14. It shows the /l/ in the sequences [film], taken from motion picture X-rays of the English word film, the German word filmen, the Spanish word filmar and the French word filmer, spoken at conversational tempo. The X-ray film was taken at 24 frames per second. At this speed the /l/-frame on which the tongue is the most fronted does not always show the instant when the tongue tip is in contact with the dento-alveolar surface; the frames we chose are those closest to a contact.

A comparison of the four profiles shows that the American /l/ differs from the others in three major ways.

a) The tip of the tongue is more retroflex; it has a concave shape not found nearly to the same extent in the target languages.

b) The root of the tongue is nearer to the wall of the pharynx, the whole tongue being more backed.

c) For American /l/ the tongue hump and the pharyngeal constriction are in the articulatory region of back vowels, perhaps nearest an [ɔ]. (However, the lip rounding and the large pharyngeal cavity shape of back vowels are not visible here; the relation between /l/ and back vowels is therefore limited.)

In the target languages, the tongue hump is in the region of front vowels, perhaps nearest an [ɛ].

The articulatory position of /l/ consonants in Figure 14 oppose English to the three target languages. It is a first

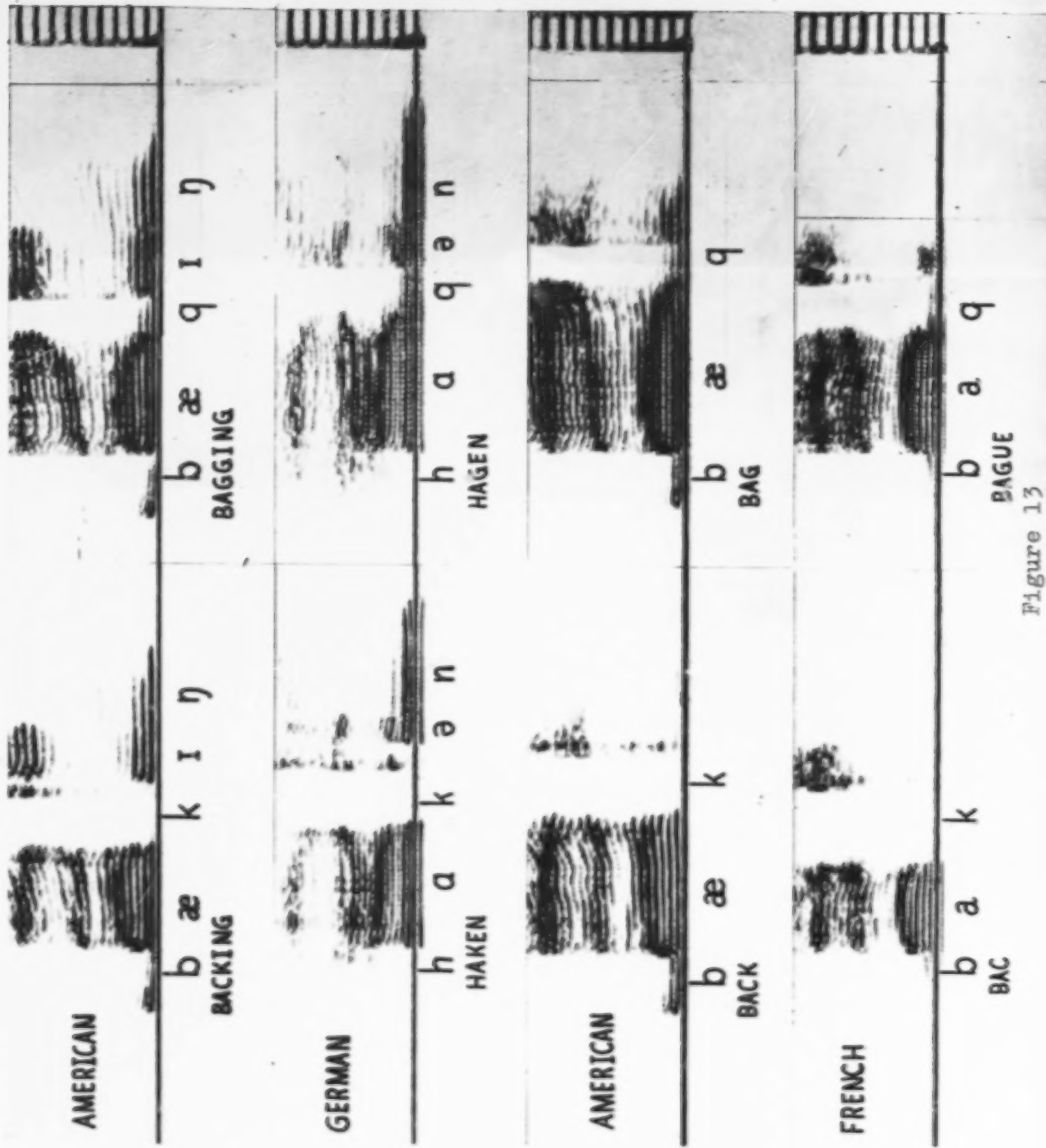


Figure 13

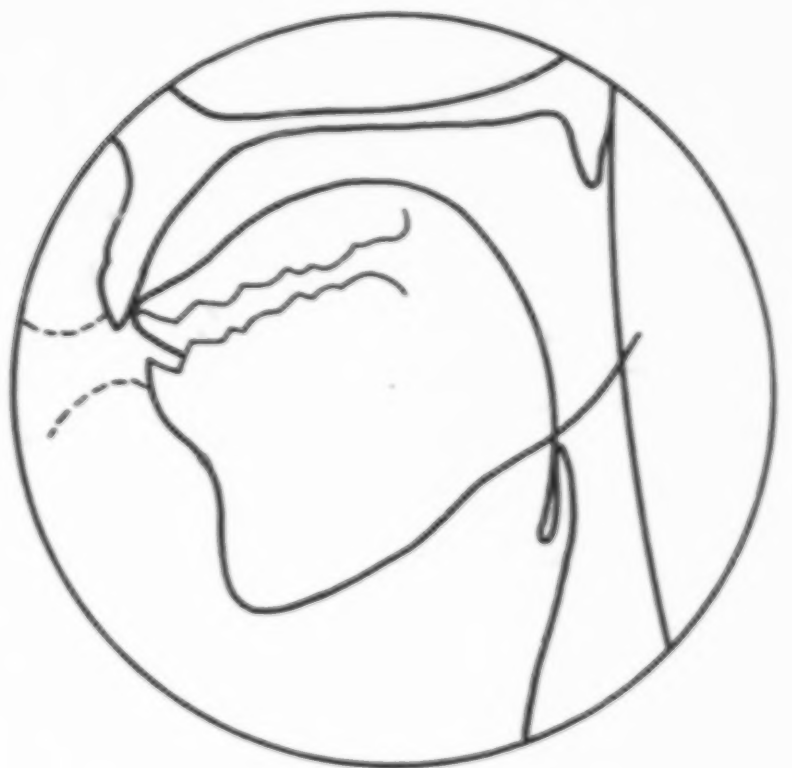
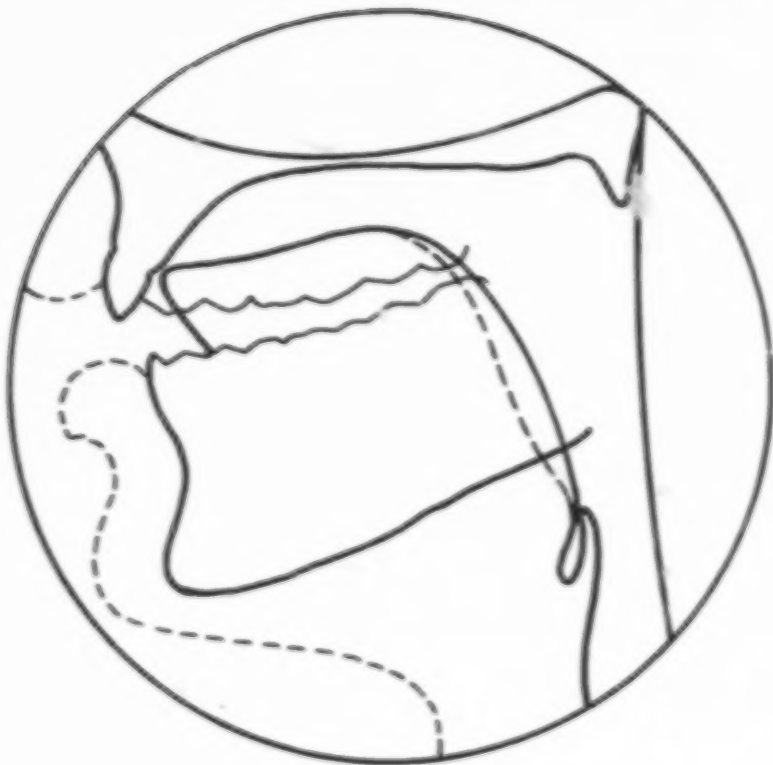
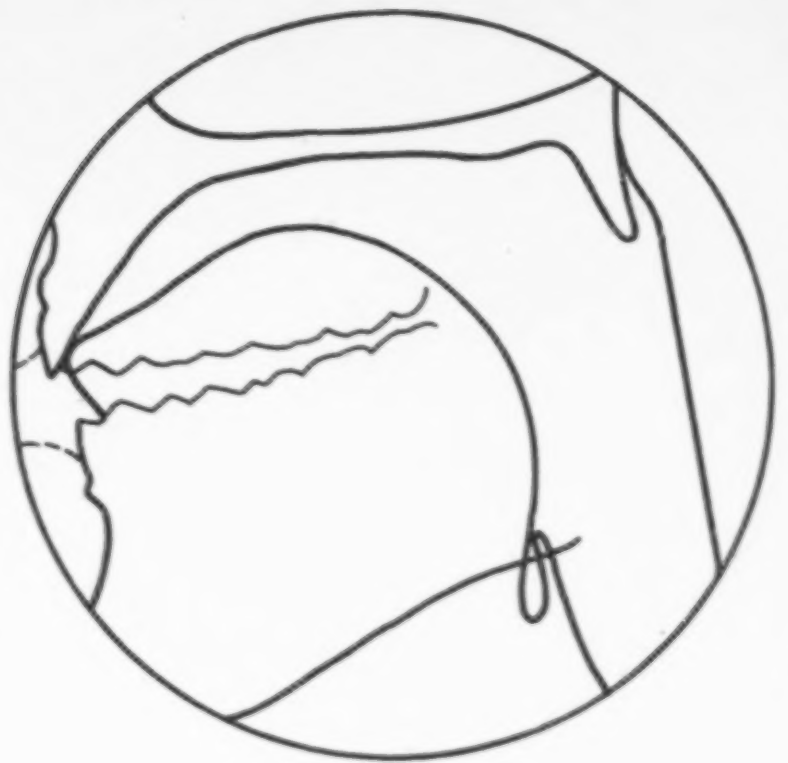


Figure 14

contribution toward defining the most characteristic tongue shape by means of motion picture X-ray.

8. CONSONANT WEAKENING

Plosive consonants tend to weaken between vowels to a degree that depends on the language. To compare the degree of weakening of lenis plosives /b/, /d/, /g/ in English and in Spanish, we are studying motion picture X-rays of comparative sentences like:

Gobierno el abogado sin duda en Vegas.

Dan was good about bidding long ago.

in which /b/, /d/, /g/ are in weak positions that facilitate comparison between the two languages, and also in strong positions for comparison within each language.

Figure 15 offers a comparison of intervocalic /b/ in motion picture sequences taken at ~~24~~⁶⁰ frames per second. It shows how the Spanish lack of complete closure of the lips contrasts with the complete closure of English (high speed cinema will show this better).

In the preceding example, Spanish is set apart from the three other languages, since German and French behave rather like English in completely closing their intervocalic lenis plosives.

In the following example, English stands apart from the three target languages. Post-stress intervocalic /t/ of American English tends to voice and to shorten its closure to the point where it is a near flap. Figure 16 offers a comparison of pre-stress initial /t/ and post-stress intervocalic /t/, articulated by the same American subject. (This was taken at 60 frames per second. Higher speed cineradiography will show this even better when it is made possible without danger of radiation.)

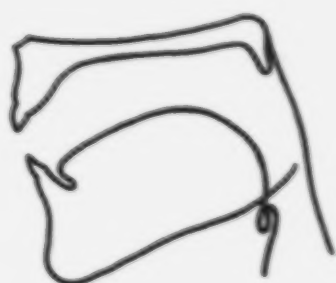
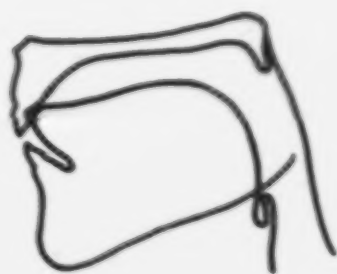
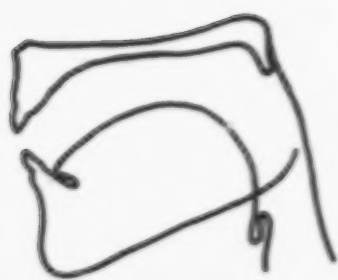


English

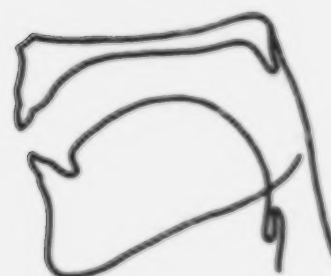
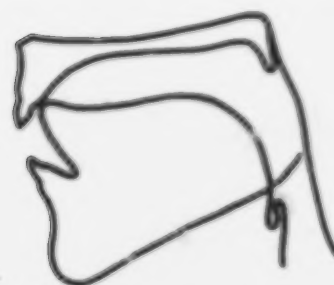
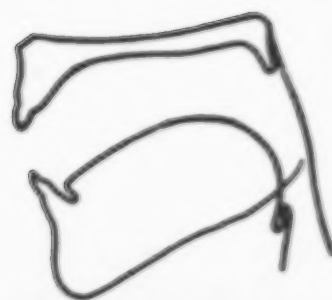


Spanish

Figure 15



Pre-stress



Post-stress

Figure 16

9. VOWEL ANTICIPATION OF CONSONANTS

There are marked differences among languages in the degree of anticipation of the articulatory position of a subsequent vowel during articulation of a consonant. This is being studied principally by the correlation of lip and tongue action where dental consonants are followed by rounded vowels, as in [du], and where labial consonants are followed by fronted vowels, as in [bi]; but it also shows in the degree of lip protrusion when a labial consonant is followed by a rounded vowel, as in [bu]. Front-view and X-ray motion pictures of sentences of the following types are being used to compare vowel anticipation during consonant articulation in English and French:

It will do to say "boo" when they coo.
 Monsieur Ledoux est dans la boue jusqu'au cou.
 McKee has a bee in his tea.
 Chez qui sont partis ses habits?

Figure 17 compares the degree of lip protrusion and roundness at the moment the tongue contacts the teeth or alveols. The lips being more protruded in French, vowel anticipation is assumed to be more pronounced.

10. SPEED OF CONSONANT ARTICULATION

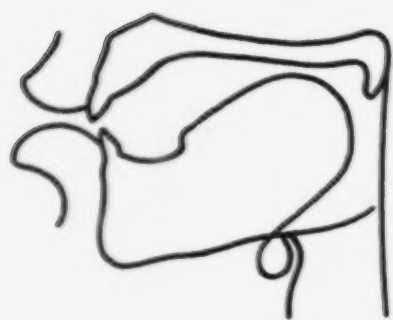
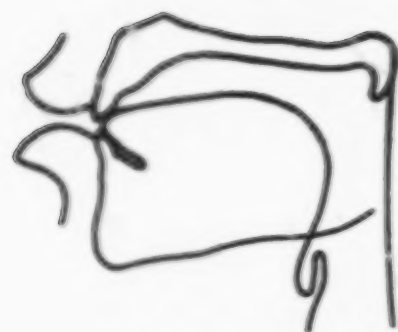
The speed of closing, and especially the speed of opening, in consonant articulation may vary considerably from one language to another. An impression of syllabic heaviness results from slow movements. To compare English and Spanish, we have made X-ray motion pictures of comparative sentences including mostly stop consonants, such as:

A publication is economical in the capital.
Una publicación es económica en la capital.

Results will not be significant until we can use high speed photography of at least five times the normal speed.



English



French

Figure 17

11. CONSONANT FRONTING

The most productive articulator is the tongue tip; it is used for over half of all the consonants in English: /θ, ð, s, z, ʃ, ʒ, tʃ, dʒ, t, d, n, l, r/. How far forward the tongue tip advances and how it is shaped in contact with teeth, alveols, or palate is different in each language. The point of articulation of apical consonants (all those that use the tip of the tongue) is therefore very important in creating the characteristic auditory impression of a language. It contributes to the degree of what is known as "frontal resonance." The apical consonants /t, d, n, l, s/ are being compared in English and Spanish in motion picture X-rays of ~~the~~ sentences like:

Notify Tony and Lois Taylor to save and donate a day late or so.
Todos los setenta fines no son deleitosos.

The apical consonants /t, d, n, l, s, z, ʃ/ before back vowels are being compared in English and French in motion picture X-rays of:

The low dome shone in sober and noble tones.
Il ose aussi verser un seau d'eau chaude sur nos autos jaunes.

Before front vowels the comparative sentences are:

Take the same native shale from the lake in the dale.
Les aînés ont aidé chez les Leger l'été passé.

In Figures 18 and 19, examples of fronting in English, as compared with the target languages, offer tongue profiles of /l/ consonants adjacent to comparable front vowels, taken from motion picture X-rays at 24 frames per second. The native subjects of Figure 19 are all different from those of Figure 18. Even though the selected frames were the ones with maximal fronting, few of them show an actual contact of the tongue. This can be expected at a speed of only 24 frames per second. In both figures, a more pronounced fronting of the target languages than of English can be observed not only at the tip of the tongue but also in back

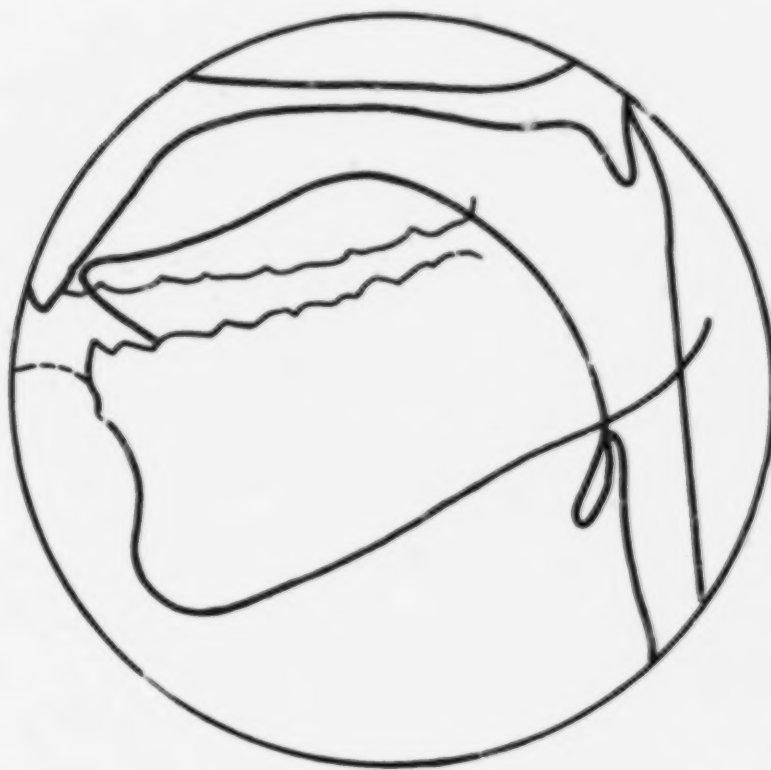
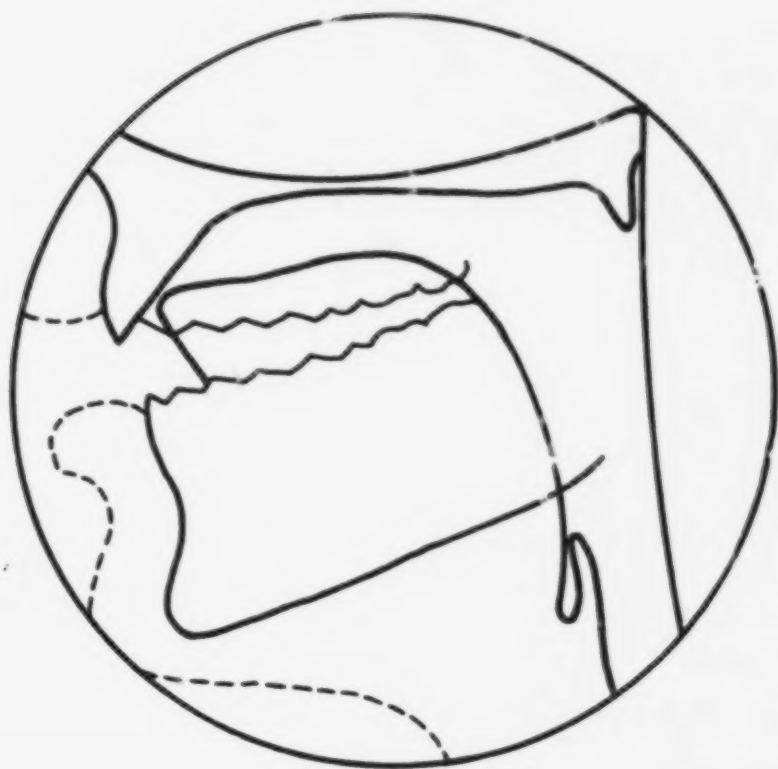
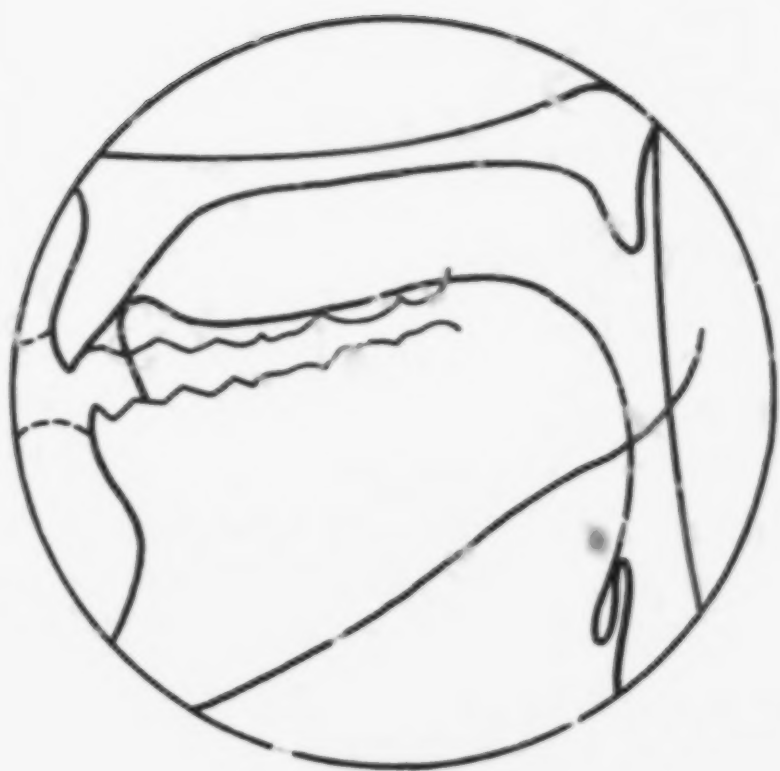
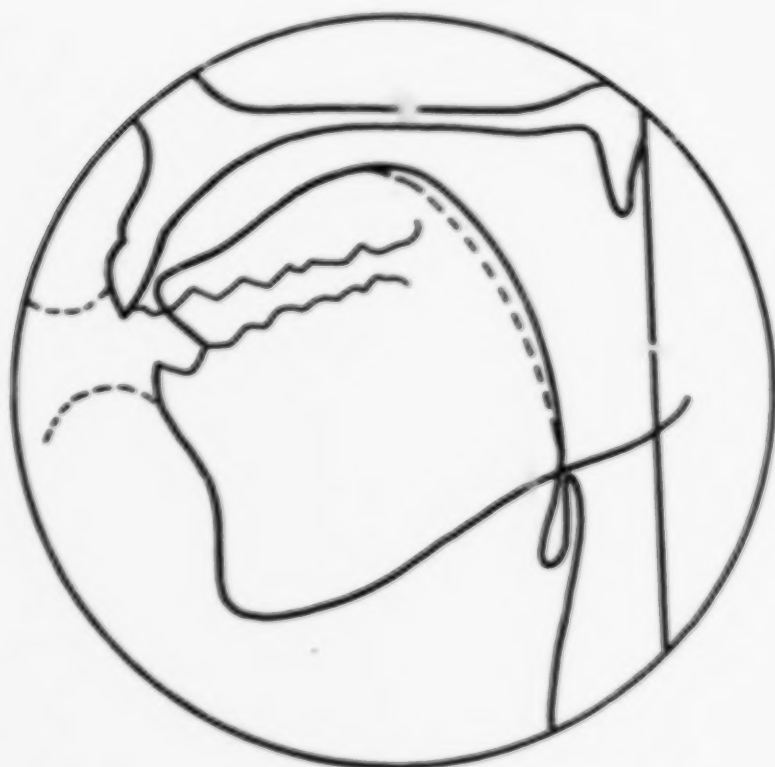
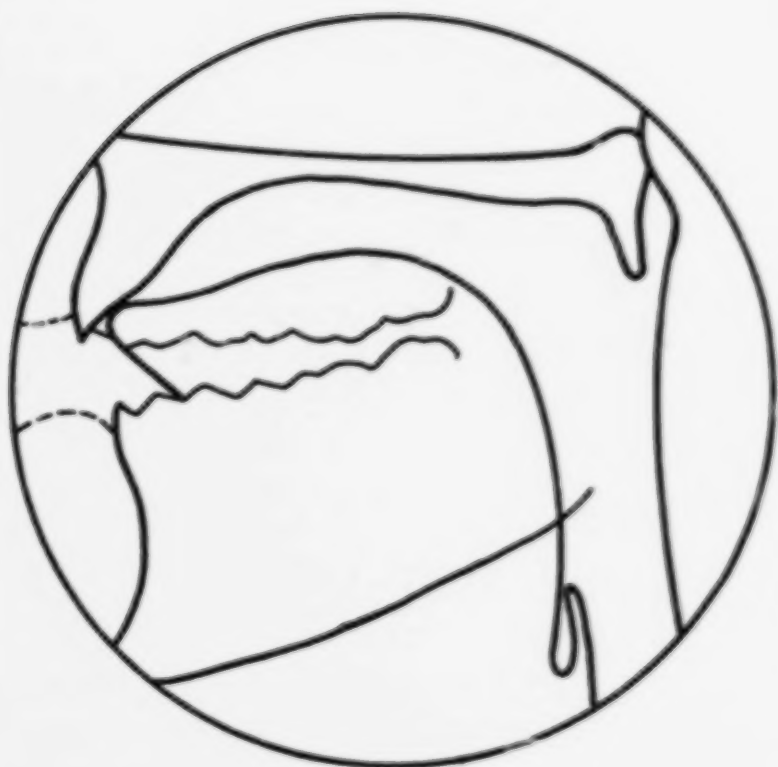
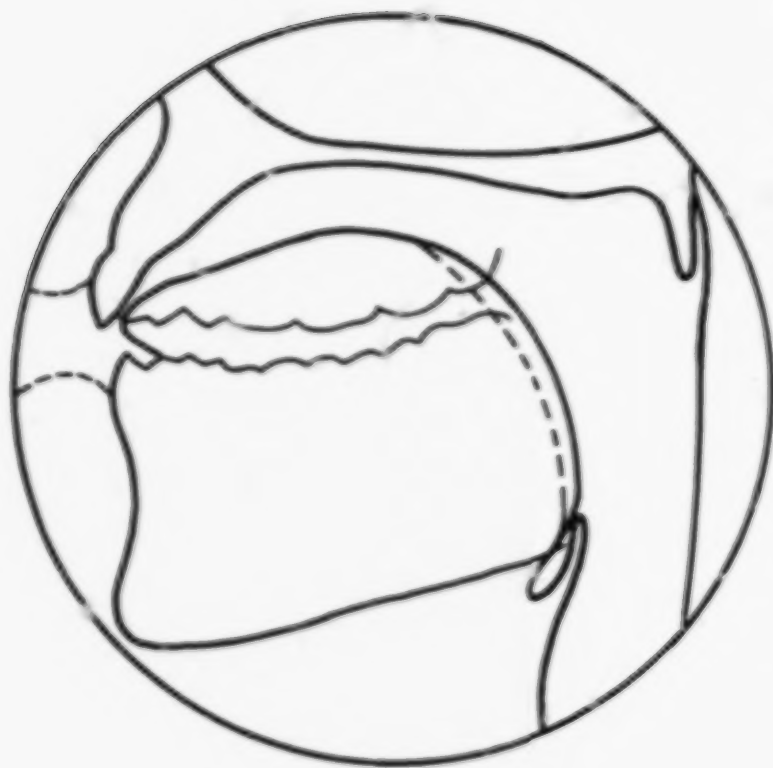
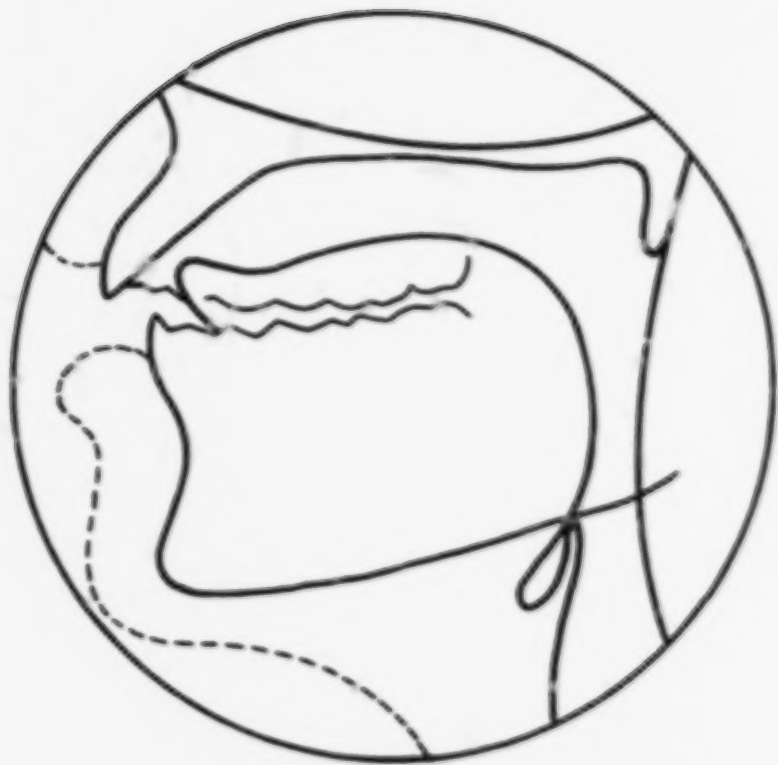


Figure 18

Figure 19



where the distance between the root of the tongue and the pharyngeal wall is considerably greater.

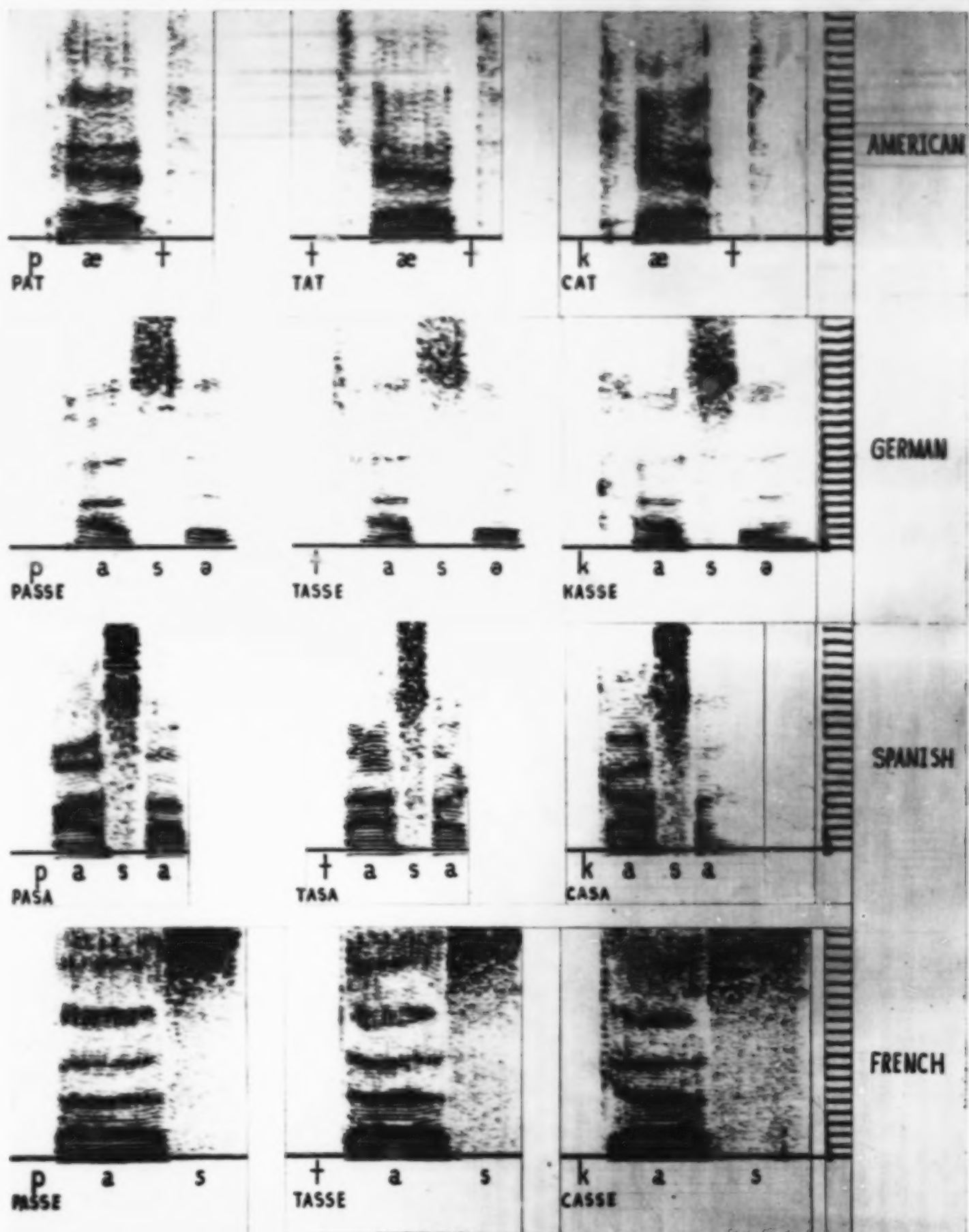
12. CONSONANT ASPIRATION

In the articulation of fortis plosive consonants, the production of an h-like sound between the separation of the organs (the explosion) and the voiced beginning of the following vowel seems to depend on how long the glottis remains open -- or how late the glottis is in closing -- after the vocal tract has opened.

Experiments by artificial speech synthesis have shown that aspiration, having no first formant but only second and above, is more like /h/ than like a whispered vowel, which has the first formant as well as the others. In fact, aspiration is characterized on spectrograms as noise (turbulent sound) in the second and third formants. It starts with the burst of noise of the explosion, and ceases when harmonics appear in the three lower formants. Aspiration appears in English and German fortis plosives, not in the Spanish and French ones. It plays an important part in characterizing the auditory impression of English and German, but it is never distinctive. (For its place of occurrence, see the chapter on Distribution.)

Figure 20 offers spectrographic examples of fortis consonants in our four languages. It shows the typical extent of aspiration in English and German, and the practical absence of it in Spanish and French. In the latter two languages, the glottis closes to vibrate for the following vowel at the instant the vocal tract opens for the explosion of the plosive.

Figure 20



13. CONSONANT AFFRICATION

Affrication creates a problem when American students learn German, Spanish or French. It is the friction noise that is added to plosive consonants, such as p, t, k, b, d, g, when the separation of the articulators for the explosion is too slow. This friction noise is evidently produced at the same place of articulation as the pure consonant; when affricated, [pa], [ta] become [pfa], [tsa], respectively. Distinctive affricates are found in English: cheap /tʃip/, jeep /dʒip/; in German: Pfund /pfunt/, zenn /tsen/; Spanish cnico /tʃiko/; but not in French.

Non-distinctive affrication most frequently occurs when English speakers attempt to pronounce dental plosive consonants before fronted vowels. A comparison of syllables like [ti], [ty] pronounced by speakers of English, French, Spanish and German before the X-ray camera will show the extent of English affrication. On spectrograms, affrication must not be confused with aspiration. The latter appears as noise in the second and third formants, the former as a separate slice of non-periodic sound, independent of the formants.

14. CONSONANT PALATALIZATION

Palatalization, which also creates a problem when American students learn German, Spanish or French, is the tendency to combine a consonant articulation with the yod that immediately follows rather than keeping them separate. The result of palatalization is a single consonant whose manner of articulation is that of the first consonant, and whose place of articulation is that of a yod, or close to it. Spanish /ɲ/ and /j/ (calle, baño), as well as French /j/ (montagne) are good examples of palatalized consonants. They are inherited from a time when Spanish and French did palatalize -- today, these languages do not palatalize, any more than German does. The English tendency to palatalize is still very much alive, and can be heard in utterances like

miss you, bet you, bid you, which are frequently pronounced [mɪʃu], [betʃu], [bɪdʒu]. When an American student is given for the first time the stimulus "monsieur" [mɛsjø], he invariably repeats [mɛʃe] because he has the habit of combining s and yod when they appear in the same syllable. If he does not combine them, as in association [əsosi'eʃən], it is because he feels that they belong to two separate syllables. These phenomena of palatalization and absence of it are being studied by spectrography and cineradiography.

15. FINAL CONSONANT RELEASE

Final consonants are released much more frequently in French than in the other languages. The reopening of the vocal tract after closing for the final consonant produces a most characteristic auditory effect: a syllabic embryo occurs which is generally voiced after lenis consonants and whispered after fortis consonants. A non-distinctive feature, it is nevertheless important. Teaching American students of French to make this release correctly and regularly, especially after n, m and l, is not a small matter.

A statistical comparison of the occurrence of final release in the four languages is underway and the articulatory comparison of its manner and extent will be done by cineradiography and spectrography.

16. CONSONANT VOICING

Literally, voiced means that the vocal cords vibrate, and voiceless that they do not. Even though the vocal cord action does not always contribute to distinguishing den from ten, seize from cease in English, the traditional terms voiced, voiceless, voicing will be retained here with the qualification that they correspond to the den/ten type of contrast, whatever the acoustic

factors may actually be. The meaning of voiced, then, will not be limited to vocal cord vibration, or presence of the fundamental frequency; it will be (arbitrarily) the name of the complete distinctive feature by which den is distinguished from ten.

The components of voicing are numerous, have varied degrees of power and differ from language to language. We have investigated them for English by spectrographic synthesis and will use the same technique for the other languages. By manipulating artificial speech patterns on our speech synthesizers, it was possible to isolate the acoustic correlates of voicing and to test the relative power of each one by ear.

Figure 21 shows how the stimulus of a single acoustic correlate at a time can be observed. The upper row offers a series of six spectrographic patterns in which the duration of the s/z noise (the dotted rectangles) is varied in steps of three centiseconds from 18 cs to 3 cs, while the formant pattern is fixed. When these patterns are passed through the speech synthesizer, one hears a gradual change from race to raise. In the middle row the noise is fixed at a medium duration and the length of the pattern is the variable. Again when this series is passed through the machine the sound gradually changes from race to raise. In the bottom row, the two factors are combined in reasonable proportions. Race is produced by shorter formants and longer noise, raise by longer formants and shorter noise.

Figure 22 summarizes the isolation of six factors of voicing for the fricatives of English, each one contributing to the racing/raising distinction. Patterns A and B oppose a long noise to a short one: racing becomes raising. Patterns C and D oppose the absence of fundamental frequency to its presence: racing does not change sufficiently to be heard as raising. Patterns D and E oppose the absence of low intensity formants during the noise (called closure formants) to their presence: racing clearly becomes raising. Patterns F and G oppose a high intensity in the s noise to a low intensity: the change from racing to raising is not convincing. Patterns H and B oppose a cutback in the first formant to its absence: racing becomes raising. Patterns

Figure 21

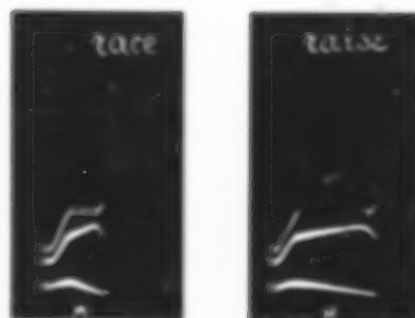
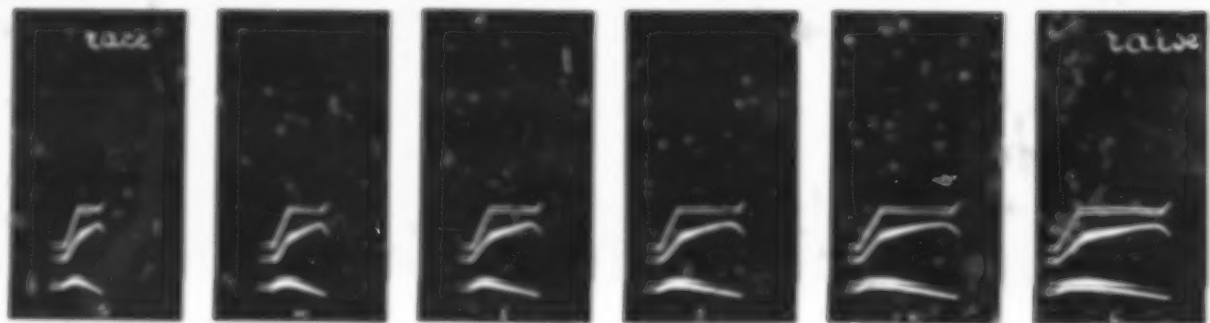
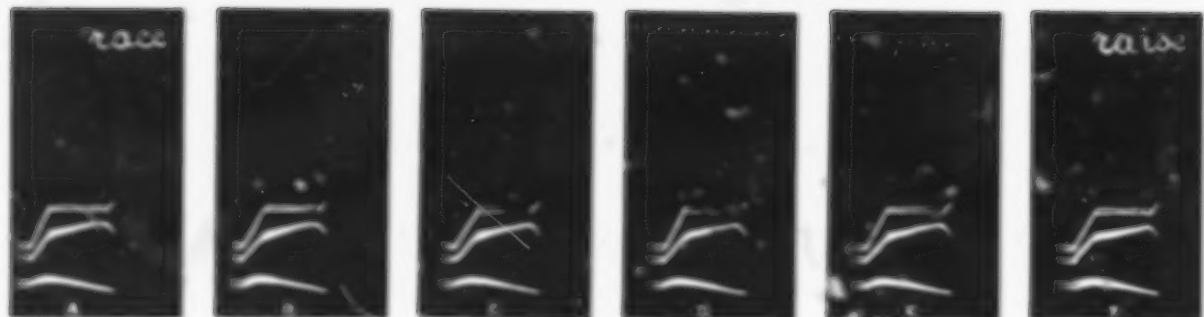


Figure 22



I and J oppose short formants before the noise to long ones: racing changes to raising.

Figure 23 summarizes the isolation of six factors of voicing for the plosive consonants in English. The pattern of the word gap is modified to make it sound like cap or gab. Patterns A and B are opposed by a first formant cutback in B: gap becomes cap. Patterns A and C are opposed by aspiration in C -- the three formants have noise instead of harmonics for about five centiseconds: gap becomes cap. Patterns A and D are opposed by a combination of aspiration and cutback in D: gap becomes cap. Patterns E and F are opposed by an intense burst of noise in F: there is a slight change from gap to cap. Patterns F and G are opposed by a fundamental during closure in G: cap does not clearly change to gap; it is still heard as cap by most listeners. Patterns H and I are opposed by a longer silence in H before the final release (this silence in plosives corresponds to the noise in fricatives): gap becomes gab. Patterns I and K are opposed by a shorter duration of the formants in I: gap becomes gab.

Five of the six factors of voicing presented here also exist for fricatives. The one that is not shared by fricatives is aspiration. And the one that exists in the fricatives but is not shared by the plosives is the low intensity closure vowel. Altogether, seven factors of voicing have been isolated for English. Preliminary testing indicates that the strongest of the seven factors are the aspiration and the closure formants. The weakest are the fundamental during closure and the intensity of the noise. The other three -- duration of silence, duration of the formants, and first formant cutback, have medium power.

We may conclude that in English plosives the greatest contribution to voicing is made by the absence of aspiration, and the weakest by the presence of a fundamental -- the very factor that gives the name to the complete distinctive feature and the only one named to explain voicing in most manuals of phonetics.

The same type of investigation for the plosives and fricatives of German, Spanish and French will, it is hoped, disclose

Figure 23



the acoustic correlates of voicing and their relative power in those languages.

Having discovered the acoustic correlates of voicing by electronic techniques, we can now list the articulatory correlates which correspond to each of the acoustic ones.

1. Aspiration is produced by keeping the glottis open at the time the organs of articulation separate and for a few more centiseconds, before the vocal cords join to vibrate.

2. Closure formants indicate that, during the friction noise of a voiced fricative, there is resonance in the whole vocal tract, and not only in the portion that is in front of the constriction.

3. Duration of noise (in fricatives) or silence (in plosives) is produced by maintaining the organs constricted for an accountable length of time.

4. Duration of formants is produced by maintaining the glottal source of sound in action for an accountable length of time while the cavities resonate that sound.

5. First formant cutback indicates that just before and just after the consonant constriction the pharyngeal (back) cavity of the vocal tract does not resonate efficiently.

6. Noise intensity is produced by building up pressure behind the constriction before allowing the opening which releases the formants. Pressure built up in the mouth depends on air pressure from the lungs and organic pressure at the constriction. The more intense the air pressure, the stronger the organic pressure must be to stop it.

7. The fundamental frequency, or first harmonic, is produced by the periodic vibration of the vocal cords.

In short, voicing, one of the most common distinctive features of speech, is also one of the most complex. It can be dissected into at least seven acoustic or articulatory features of speech, which are not all of equal importance for the identification of consonants.